

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 18:02:37 ; Search time 23 seconds  
(without alignments)  
2666.594 Million cell updates/sec

Title: US-09-980-403-2  
Perfect score: 1188  
Sequence: 1 MDLPRGLVAVALLSLWPGET.....FRSARRRRFGDTPTEKVLK 1188

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| 1          | 21    | 1.8         | 158    | 4  | US-09-312-283C-391   |
| 2          | 17    | 1.4         | 85     | 4  | US-09-312-283C-390   |
| 3          | 15    | 1.3         | 214    | 4  | US-10-061-658-5      |
| 4          | 15    | 1.3         | 214    | 4  | US-10-061-658-6      |
| 5          | 15    | 1.3         | 214    | 4  | US-10-061-658-9      |
| 6          | 14    | 1.2         | 1183   | 4  | US-09-532-310B-5     |
| 7          | 11    | 0.9         | 122    | 6  | 5196511-9            |
| 8          | 11    | 0.9         | 1183   | 4  | US-09-532-310B-6     |
| 9          | 9     | 0.8         | 9      | 1  | US-08-462-128-34     |
| 10         | 9     | 0.8         | 9      | 1  | US-08-463-180-34     |
| 11         | 8     | 0.7         | 17     | 1  | US-08-173-497-20     |
| 12         | 8     | 0.7         | 17     | 1  | US-08-236-889-20     |
| 13         | 8     | 0.7         | 17     | 1  | US-08-485-618-20     |
| 14         | 8     | 0.7         | 17     | 1  | US-08-382-652-20     |
| 15         | 8     | 0.7         | 17     | 2  | US-08-605-672-20     |
| 16         | 8     | 0.7         | 17     | 2  | US-08-482-293A-20    |
| 17         | 8     | 0.7         | 17     | 2  | US-08-943-363-20     |
| 18         | 8     | 0.7         | 17     | 3  | US-09-193-043-20     |
| 19         | 8     | 0.7         | 17     | 4  | US-09-688-307A-20    |
| 20         | 8     | 0.7         | 17     | 4  | US-09-350-259-20     |
| 21         | 8     | 0.7         | 76     | 1  | US-08-848-652-4      |
| 22         | 8     | 0.7         | 76     | 3  | US-09-083-521-5      |
| 23         | 8     | 0.7         | 80     | 1  | US-08-848-252-2      |
| 24         | 8     | 0.7         | 95     | 1  | US-07-626-618A-2     |
| 25         | 8     | 0.7         | 95     | 1  | US-07-928-611-2      |
| 26         | 8     | 0.7         | 95     | 1  | US-08-333-977-2      |
| 27         | 8     | 0.7         | 95     | 2  | US-08-487-811A-2     |
| 28         | 8     | 0.7         | 95     | 2  | US-09-060-694-2      |
| 29         | 8     | 0.7         | 95     | 2  | US-09-378-074-2      |
| 30         | 8     | 0.7         | 95     | 5  | PCT-US93-07370-2     |
| 31         | 8     | 0.7         | 97     | 4  | US-08-973-544-4      |
| 32         | 8     | 0.7         | 104    | 4  | US-08-973-544-3      |
| 33         | 8     | 0.7         | 112    | 4  | US-09-489-039A-10236 |
| 34         | 8     | 0.7         | 179    | 4  | US-09-419-568F-28    |
| 35         | 8     | 0.7         | 179    | 4  | US-09-354-243B-28    |
| 36         | 8     | 0.7         | 179    | 4  | US-09-870-574-2      |
| 37         | 8     | 0.7         | 183    | 4  | US-09-252-991A-25905 |
| 38         | 8     | 0.7         | 234    | 3  | US-09-040-483-5      |
| 39         | 8     | 0.7         | 234    | 4  | US-08-740-036-5      |
| 40         | 8     | 0.7         | 315    | 1  | US-08-118-270-28     |
| 41         | 8     | 0.7         | 315    | 5  | PCT-US93-08528-48    |
| 42         | 8     | 0.7         | 345    | 4  | US-09-489-039A-7417  |
| 43         | 8     | 0.7         | 374    | 3  | US-09-046-736-4      |
| 44         | 8     | 0.7         | 387    | 1  | US-07-626-618A-17    |
| 45         | 8     | 0.7         | 387    | 1  | US-08-086-438C-3     |
| 46         | 8     | 0.7         | 387    | 1  | US-08-056-051-2      |
| 47         | 8     | 0.7         | 387    | 1  | US-07-928-611-18     |
| 48         | 8     | 0.7         | 387    | 1  | US-08-333-977-17     |
| 49         | 8     | 0.7         | 387    | 1  | US-08-449-153-1      |
| 50         | 8     | 0.7         | 387    | 1  | US-08-434-877-3      |
| 51         | 8     | 0.7         | 387    | 1  | US-08-449-651-1      |
| 52         | 8     | 0.7         | 387    | 2  | US-08-487-811A-18    |
| 53         | 8     | 0.7         | 387    | 3  | US-08-833-703-1      |
| 54         | 8     | 0.7         | 387    | 3  | US-08-475-742-4      |
| 55         | 8     | 0.7         | 387    | 3  | US-09-060-694-18     |
| 56         | 8     | 0.7         | 387    | 4  | US-09-378-074-18     |
| 57         | 8     | 0.7         | 387    | 4  | US-08-261-293-4      |
| 58         | 8     | 0.7         | 387    | 5  | PCT-US93-07370-18    |
| 59         | 8     | 0.7         | 419    | 1  | US-08-056-051-4      |
| 60         | 8     | 0.7         | 419    | 1  | US-07-928-611-20     |
| 61         | 8     | 0.7         | 419    | 2  | US-08-487-811A-20    |
| 62         | 8     | 0.7         | 419    | 3  | US-09-060-694-20     |
| 63         | 8     | 0.7         | 419    | 4  | US-09-378-074-20     |
| 64         | 8     | 0.7         | 419    | 5  | PCT-US93-07370-20    |
| 65         | 8     | 0.7         | 453    | 4  | US-09-252-991A-17394 |
| 66         | 8     | 0.7         | 467    | 1  | US-08-056-051-6      |
| 67         | 8     | 0.7         | 467    | 1  | US-07-928-611-22     |
| 68         | 8     | 0.7         | 467    | 2  | US-08-487-811A-22    |
| 69         | 8     | 0.7         | 467    | 3  | US-09-046-736-2      |
| 70         | 8     | 0.7         | 467    | 3  | US-09-060-694-22     |
| 71         | 8     | 0.7         | 467    | 4  | US-09-378-074-22     |
| 72         | 8     | 0.7         | 467    | 5  | PCT-US93-07370-22    |
| 73         | 8     | 0.7         | 494    | 1  | US-08-485-618-103    |
| 74         | 8     | 0.7         | 494    | 2  | US-08-605-672-103    |
| 75         | 8     | 0.7         | 494    | 2  | US-08-482-293A-103   |
| 76         | 8     | 0.7         | 494    | 2  | US-08-943-363-103    |
| 77         | 8     | 0.7         | 494    | 3  | US-09-193-043-103    |
| 78         | 8     | 0.7         | 494    | 4  | US-09-688-307A-103   |
| 79         | 8     | 0.7         | 494    | 4  | US-09-350-259-103    |
| 80         | 8     | 0.7         | 605    | 2  | US-08-472-666-1      |
| 81         | 8     | 0.7         | 605    | 5  | PCT-US96-07615-1     |
| 82         | 8     | 0.7         | 628    | 4  | US-09-252-991A-3727  |
| 83         | 8     | 0.7         | 877    | 4  | US-09-252-991A-25547 |
| 84         | 8     | 0.7         | 1151   | 1  | US-08-286-889-37     |
| 85         | 8     | 0.7         | 1151   | 1  | US-08-485-618-37     |
| 86         | 8     | 0.7         | 1151   | 2  | US-08-362-652-37     |
| 87         | 8     | 0.7         | 1151   | 2  | US-08-605-672-37     |
| 88         | 8     | 0.7         | 1151   | 2  | US-08-482-293A-37    |
| 89         | 8     | 0.7         | 1151   | 2  | US-08-943-363-37     |
| 90         | 8     | 0.7         | 1151   | 3  | US-09-193-043-37     |
| 91         | 8     | 0.7         | 1151   | 4  | US-09-688-307A-37    |
| 92         | 8     | 0.7         | 1151   | 4  | US-09-350-259-37     |
| 93         | 8     | 0.7         | 1152   | 4  | US-08-476-062A-43    |
| 94         | 8     | 0.7         | 1152   | 5  | PCT-US96-01314-43    |
| 95         | 8     | 0.7         | 1152   | 6  | 5424399-2            |
| 96         | 8     | 0.7         | 1153   | 1  | US-08-173-497-3      |
| 97         | 8     | 0.7         | 1153   | 1  | US-08-286-889-3      |
| 98         | 8     | 0.7         | 1153   | 1  | US-08-485-618-3      |
| 99         | 8     | 0.7         | 1153   | 1  | US-08-362-652-3      |
| 100        | 8     | 0.7         | 1153   | 2  | US-08-605-672-3      |

Patent No. 5424399

## ALIGNMENTS

```
RESULT 1
US-09-312-283C-391
; Sequence 391, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions isolated from skin cells
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 391
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-391

Query Match 1.8%; Score 21; DB 4; Length 158;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 RRVQINASHLFEPKINIFH 656
DB 57 RRVQINASHLFEPKINIFH 77

RESULT 2
US-09-312-283C-390
; Sequence 390, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions isolated from skin cells
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 390
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-390

Query Match 1.4%; Score 17; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 NGTLKDSHSYQNARTGS 544
DB 28 NGTLKDSHSYQNARTGS 44

RESULT 3
US-10-061-658-5
; Sequence 5, Application US/10061658
; Patent No. 6652856
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Kotliansky, Victor
; TITLE OF INVENTION: Method for the treatment of fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/061,658
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Rat
US-10-061-658-5

Query Match 1.3%; Score 15; DB 4; Length 214;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
DB 22 DIVIVLDGNSIYPW 36

RESULT 4
US-10-061-658-6
; Sequence 6, Application US/10061658
; Patent No. 6652856
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Kotliansky, Victor
; TITLE OF INVENTION: Method for the treatment of fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/061,658
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-061-658-6

Query Match 1.3%; Score 15; DB 4; Length 214;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
DB 22 DIVIVLDGNSIYPW 36

RESULT 5
US-10-061-658-9
; Sequence 9, Application US/10061658
; Patent No. 6652856
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Kotliansky, Victor
; TITLE OF INVENTION: Method for the treatment of fibrosis
```



FILE REFERENCE: A073US  
CURRENT APPLICATION NUMBER: US/10/061,658  
CURRENT FILING DATE: 2002-02-01  
PRIOR APPLICATION NUMBER: 60/137,214  
PRIOR FILING DATE: 1999-06-01  
PRIOR APPLICATION NUMBER: 60/130,847  
PRIOR FILING DATE: 1999-04-22  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 214  
TYPE: PRT  
ORGANISM: Homo sapien  
US-10-061-658-9

Query Match 1.3%; Score 15; DB 4; Length 214;  
Best Local Similarity 100.0%; Pred. No. 6.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178  
Db 22 DIVVLGSGNSIYPW 36

## RESULT 6

US-09-532-310B-5  
Sequence 5, Application US/09532310B

Patent No. 6596276

GENERAL INFORMATION:

APPLICANT: Senger, Donald R

Detmar, Michael

Claffey, Kevin P

TITLE OF INVENTION: Method for inhibiting tumor  
angiogenesis in a living subject

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: David Prashker, Esq.

STREET: P.O. Box 5387

CITY: Magnolia

STATE: Massachusetts

COUNTRY: USA

ZIP: 01930

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.40 Mb storage

COMPUTER: Dell PC

OPERATING SYSTEM: MS DOS

SOFTWARE: Microsoft Word version 97

CURRENT APPLICATION DATA: US/09/532,310B

APPLICATION NUMBER: US/09/532,310B

FILING DATE: 22-Mar-2000

CLASSIFICATION: Unknown

ATTORNEY/AGENT INFORMATION:

NAME: David Prashker, Esq.

REGISTRATION NUMBER: 29,693

REFERENCE/DOCKET NUMBER: BIS-036

TELECOMMUNICATION INFORMATION:

TELEPHONE: (978) 525-3794

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1183 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Query Match 1.2%; Score 14; DB 4; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 0.00027;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYP 177  
|||||

Db 174 DIVVLGSGNSIYP 187

## RESULT 7

5196511-9

Patent No. 5196511

APPLICANT: PLOW, EDWARD F.; D'SOUZA, STANLEY E.

GINSBERG, MARK H.

TITLE OF INVENTION: PEPTIDES AND ANTIBODIES THAT INHIBIT

INTEGRIN-LIGAND BINDING

NUMBER OF SEQUENCES: 31

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/444,777

FILING DATE: 01-DEC-1989

SEQ ID NO: 9:

LENGTH: 22

5196511-9

Query Match 0.9%; Score 11; DB 6; Length 22;

Best Local Similarity 100.0%; Pred. No. 0.0054;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAPMY 505

Db 11 TDVLLVGAPMY 21

## RESULT 8

US-09-532-310B-6

Sequence 6, Application US/09532310B

Patent No. 6596276

GENERAL INFORMATION:

APPLICANT: Senger, Donald R

Detmar, Michael

Claffey, Kevin P

TITLE OF INVENTION: Method for inhibiting tumor

angiogenesis in a living subject

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: David Prashker, Esq.

STREET: P.O. Box 5387

CITY: Magnolia

STATE: Massachusetts

COUNTRY: USA

ZIP: 01930

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.40 Mb storage

COMPUTER: Dell PC

OPERATING SYSTEM: MS DOS

SOFTWARE: Microsoft Word version 97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/532,310B

FILING DATE: 22-Mar-2000

CLASSIFICATION: Unknown

ATTORNEY/AGENT INFORMATION:

NAME: David Prashker, Esq.

REGISTRATION NUMBER: 29,693

REFERENCE/DOCKET NUMBER: BIS-036

TELECOMMUNICATION INFORMATION:

TELEPHONE: (978) 525-3794

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 1183 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-532-310B-6

Query Match 0.9%; Score 11; DB 4; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 0.19;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 495 TDVLLVGAPWY 505  
 |||||  
 Db 506 TDVLLVGAPWY 516

RESULT 9  
 US-08-462-128-34  
 ; Sequence 34, Application US/08462128  
 ; Patent No. 5686059  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Goetinck, Paul F.  
 ; APPLICANT: Tondravi, Mehرداد  
 ; TITLE OF INVENTION: CARTILAGE MATRIX PROTEIN AND METHODS FOR  
 ; TITLE OF INVENTION: USE  
 ; NUMBER OF SEQUENCES: 43  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Lahive & Cockfield  
 ; STREET: 60 State Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/462,128  
 ; FILING DATE:  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/006,096  
 ; FILING DATE: 15-JAN-1993  
 ; APPLICATION NUMBER: US 07/866,403  
 ; FILING DATE: 10-APR-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Myers, Paul L.  
 ; REGISTRATION NUMBER: 35,965  
 ; REFERENCE/DOCKET NUMBER: MGP-005  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5941  
 ; INFORMATION FOR SEQ ID NO: 34:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 9 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; US-08-462-128-34

Query Match 0.8%; Score 9; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVIVLDGS 172  
 |||||  
 Db 1 DIVIVLDGS 9

RESULT 10  
 US-08-463-180-34  
 ; Sequence 34, Application US/08463180  
 ; Patent No. 5741670  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Goetinck, Paul F.  
 ; APPLICANT: Tondravi, Mehرداد  
 ; TITLE OF INVENTION: CARTILAGE MATRIX PROTEIN AND METHODS FOR  
 ; TITLE OF INVENTION: USE  
 ; NUMBER OF SEQUENCES: 43  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Lahive & Cockfield  
 ; STREET: 60 State Street

CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/463,160  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/006,096  
 ; FILING DATE: 15-JAN-1993  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/866,403  
 ; FILING DATE: 10-APR-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Myers, Paul L.  
 ; REGISTRATION NUMBER: 35,965  
 ; REFERENCE/DOCKET NUMBER: MGP-005DV  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5941  
 ; INFORMATION FOR SEQ ID NO: 34:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 9 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; US-08-463-180-34

Query Match 0.8%; Score 9; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVIVLDGS 172  
 |||||  
 Db 1 DIVIVLDGS 9

RESULT 11  
 US-08-173-497-20  
 ; Sequence 20, Application US/08173497  
 ; Patent No. 5437958  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gallatin, W. Michael  
 ; APPLICANT: Van Der Vieren, Monica  
 ; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha  
 ; TITLE OF INVENTION: Subunit  
 ; NUMBER OF SEQUENCES: 29  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 ; STREET: 233 S. Wacker Drive, 6300 Sears Tower  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606-6402  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/173,497  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. 5437958and, Greta E.  
 ; REGISTRATION NUMBER: 35,302  
 ; REFERENCE/DOCKET NUMBER: 278666/31363

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-173-497-20

Query Match 0.7%; Score 8; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 12  
US-08-286-889-20  
Sequence 20, Application US/08286889  
Patent No. 5470953  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Mich  
TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/286,889  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32168  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-286-889-20

Query Match 0.7%; Score 8; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 13  
US-08-485-618-20  
Sequence 20, Application US/08485618  
Patent No. 5728533  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,618  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32297  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-485-618-20

Query Match 0.7%; Score 8; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 14  
US-08-362-652-20  
Sequence 20, Application US/08362652  
Patent No. 5766850  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago

; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/362,652  
; FILING DATE:  
; CLASSIFICATION: 430  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32391  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-362-652-20

Query Match 0.7%; Score 8; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 15  
US-08-605-672-20  
; Sequence 20, Application US/08605672  
; Patent No. 5817515  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/605,672  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889

; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-605-672-20

Query Match 0.7%; Score 8; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 16  
US-08-482-293A-20  
; Sequence 20, Application US/08482293A  
; Patent No. 5831029  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,293A  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-482-293A-20

Query Match 0.7%; Score 8; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 17  
US-08-943-363-20  
; Sequence 20, Application US/08943363  
; Patent No. 5837478  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 114  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/943,363  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-943-363-20

Query Match 0.7%; Score 8; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 18  
US-09-193-043-20  
; Sequence 20, Application US/09193043  
; Patent No. 6251395  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 6251395el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/193,043  
; CURRENT FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: dog  
US-09-193-043-20

Query Match 0.7%; Score 8; DB 3; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 19  
US-09-688-307A-20  
; Sequence 20, Application US/09688307A  
; Patent No. 6432404  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 6432404el Human Beta-2  
; FILE REFERENCE: 27866/36646  
; CURRENT APPLICATION NUMBER: US/09/688,307A  
; CURRENT FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 09/193,043  
; PRIOR FILING DATE: 1998-11-16  
; PRIOR APPLICATION NUMBER: 08/605,672  
; PRIOR FILING DATE: 1996-02-22  
; PRIOR APPLICATION NUMBER: 08/173,497  
; PRIOR FILING DATE: 1993-12-23  
; PRIOR APPLICATION NUMBER: 08/286,889  
; PRIOR FILING DATE: 1994-08-05  
; PRIOR APPLICATION NUMBER: 08/362,652  
; PRIOR FILING DATE: 1994-12-21  
; PRIOR APPLICATION NUMBER: 08/943,363  
; PRIOR FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: dog  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 13  
; OTHER INFORMATION: Xaa = any or unknown amino acid  
US-09-688-307A-20

Query Match 0.7%; Score 8; DB 4; Length 17;

Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
Db 1 LVVGAPLE 8

## RESULT 20

US-09-350-259-20  
; Sequence 20, Application US/09350259  
; Patent No. 6620915  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 6620915el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350,259  
; EARLIER FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: 09/193,043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: dog  
US-09-350-259-20

Query Match 0.7%; Score 8; DB 4; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
Db 1 LVVGAPLE 8

## RESULT 21

US-08-848-252-4  
; Sequence 4, Application US/08848252  
; Patent No. 5804177  
; GENERAL INFORMATION:  
; APPLICANT: Humphries, Keith R.  
; TITLE OF INVENTION: METHOD OF USING CD24 AS A CELL MARKER  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bereskin & Parr  
; STREET: 40 King Street West  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 3Y2  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/848,252  
; FILING DATE: 29-APR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/538,052  
; FILING DATE:

; APPLICATION NUMBER: US 08/151,672  
; FILING DATE: 15-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McDiarmid, Shona S.  
; REGISTRATION NUMBER: P-38,798  
; REFERENCE/DOCKET NUMBER: 3158-028  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 364-7311  
; TELEFAX: (416) 361-1398  
; TELETYPE: 06-23115  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 76 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-848-252-4

Query Match 0.7%; Score 8; DB 1; Length 76;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLALL 1159  
Db 12 GLLALL 19

## RESULT 22

US-09-083-521-5  
; Sequence 5, Application US/09083521  
; Patent No. 6048970  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/083,521  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CERRONE, MICHAEL C.  
; REGISTRATION NUMBER: 39,132  
; REFERENCE/DOCKET NUMBER: PP-0527 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 76 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 1216498  
US-09-083-521-5

Query Match 0.7%; Score 8; DB 3; Length 76;  
Best Local Similarity 100.0%; Pred. No. 11;

```
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 12 GLLLLALL 19

RESULT 23
US-08-848-252-2
; Sequence 2, Application US/08848252
; Patent No. 5804177
; GENERAL INFORMATION:
; APPLICANT: Humphries, Keith R.
; TITLE OF INVENTION: METHOD OF USING CD24 AS A CELL MARKER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/848,252
; FILING DATE: 29-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/538,052
; FILING DATE:
; APPLICATION NUMBER: US 08/151,672
; FILING DATE: 15-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McDiarmid, Shona S.
; REGISTRATION NUMBER: P-38,798
; REFERENCE/DOCKET NUMBER: 3158-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; TELEX: 06-23315
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-848-252-2

Query Match 0.7%; Score 8; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 12 GLLLLALL 19

RESULT 24
US-07-626-618A-2
; Sequence 2, Application US/07626618A
; Patent No. 5422265
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5422265el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
```

```
QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 25
US-07-928-611-2
; Sequence 2, Application US/07928611
; Patent No. 5559601
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/928,611
; FILING DATE: 19920810
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 5569601nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-626-618A-2

Query Match 0.7%; Score 8; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88
```

```
;
; LENGTH: 95 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-928-611-2
Query Match 0.7%; Score 8; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 26
US-08-333-977-2
; Sequence 2, Application US/08333977
; Patent No. 5594108
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5594108el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,977
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/626,618
; FILING DATE: 7 DEC 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5594108nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 95 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-333-977-2

Query Match 0.7%; Score 8; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 27
US-08-487-811A-2
; Sequence 2, Application US/08487811A
; Patent No. 5883226
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,694
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6203998nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-487-811A-2

Query Match 0.7%; Score 8; DB 2; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 28
US-09-060-694-2
; Sequence 2, Application US/09060694
; Patent No. 6203998
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 6203998el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,694
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6203998nan, Kevin E
; REGISTRATION NUMBER: 35,303
```



REFERENCE/DOCKET NUMBER: 90,1092-MM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 95 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-060-694-2

Query Match 0.7%; Score 8; DB 3; Length 95;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 29

US-09-378-074-2  
Sequence 2, Application US/09378074  
Patent No. 6437114  
GENERAL INFORMATION:  
APPLICANT: Van Tol, Hubert H.M.  
Civelli, Olivier  
TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Allegretti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/378,074  
FILING DATE: 20-Aug-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/928,611  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6437114nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1234  
TELEFAX: 312-715-1234  
TELEX: 810-221-8317  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 95 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-378-074-2

Query Match 0.7%; Score 8; DB 4; Length 95;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

DB 81 LLLALLVL 88

RESULT 30

PCT-US93-07370-2  
Sequence 2, Application PC/TUS9307370  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/07370  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 95 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US93-07370-2

Query Match 0.7%; Score 8; DB 5; Length 95;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 31

US-08-973-544-4  
Sequence 4, Application US/08973544  
Patent No. 6338950  
GENERAL INFORMATION:  
APPLICANT: WEISS, Elisabeth  
TITLE OF INVENTION: NEW IMMUNOREGULATORY PROTEIN LST-1  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM LLP  
STREET: 655 Fifteenth St., NW, Suite 300, G St. Lobby  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/973,544  
FILING DATE: 18-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT EP 96/02663  
FILING DATE: 20-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 95109511.6  
FILING DATE: 20-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 95112201.9  
FILING DATE: 03-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitts, Monica Chin  
REGISTRATION NUMBER: 36,105  
REFERENCE/DOCKET NUMBER: P8341-7073  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-5000

TELEFAX: (202) 638-4810  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 97 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-973-544-4

Query Match 0.7%; Score 8; DB 4; Length 97;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGILLLA 1157  
16 LGGILLLA 23

Db

RESULT 32  
US-08-973-544-3  
Sequence 3, Application US/08973544  
Patent No. 6338950  
GENERAL INFORMATION:  
APPLICANT: WEISS, Elisabeth  
TITLE OF INVENTION: NEW IMMUNOREGULATORY PROTEIN LST-1  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIKAI DO, MARMEI STEIN, MURRAY & ORAM LLP  
STREET: 655 Fifteenth St., NW, Suite 300, G St. Lobby  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/973,544  
FILING DATE: 18-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT EP 96/02663  
FILING DATE: 20-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 95109511.6  
FILING DATE: 20-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 95112201.9  
FILING DATE: 03-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitts, Monica Chin  
REGISTRATION NUMBER: 36,105  
REFERENCE/DOCKET NUMBER: P8341-7073  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-5000  
TELEFAX: (202) 638-4810  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 104 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-973-544-3

Query Match 0.7%; Score 8; DB 4; Length 104;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGILLLA 1157  
16 LGGILLLA 23

Db

RESULT 33  
US-09-489-039A-10236  
Sequence 10236, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709-2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 10236  
LENGTH: 112  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-10236

Query Match 0.7%; Score 8; DB 4; Length 112;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160  
13 LLLALLLV 20

Db

RESULT 34  
US-09-419-568F-28  
Sequence 28, Application US/09419568F  
Patent No. 6331613  
GENERAL INFORMATION:  
APPLICANT: Dumoutier, Laure  
APPLICANT: Louhed, Jamila  
APPLICANT: Renauld, Jean-Christophe  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible  
FILE REFERENCE: LUD 5543.2  
CURRENT APPLICATION NUMBER: US/09/419,568F  
CURRENT FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: US09/354,243  
PRIOR FILING DATE: 1999-07-16  
PRIOR APPLICATION NUMBER: US09/178,973  
PRIOR FILING DATE: 1998-10-26  
NUMBER OF SEQ ID NOS: 29  
SEQ ID NO 28  
LENGTH: 179  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
US-09-419-568F-28

Query Match 0.7%; Score 8; DB 4; Length 179;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160  
21 LLLALLLV 28

Db

RESULT 35  
US-09-354-243B-28  
Sequence 28, Application US/09354243B  
Patent No. 6359117  
GENERAL INFORMATION:  
APPLICANT: Dumoutier, Laure  
APPLICANT: Louhed, Jamila

```

; APPLICANT: Renault, Jean-Christophe
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fa
; TITLE OF INVENTION: (TIFFS)
; TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof
; FILE REFERENCE: LUD 5543.1
; CURRENT APPLICATION NUMBER: US/09/354,243B
; CURRENT FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: US09/178,973
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 29
; SEQ ID NO 28
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-354-243B-28

Query Match      0.7%; Score 8; DB 4; Length 179;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160
DB 21 LLLLLLV 28

RESULT 36
US-09-870-574-2
; Sequence 2, Application US/09870574
; Patent No. 6551799
; GENERAL INFORMATION:
; APPLICANT: Gurney, Austin L.
; APPLICANT: Aggarwal, Sudeepa
; APPLICANT: Xie, Ming-Hong
; APPLICANT: Maruoka, Ellen M.
; APPLICANT: Foster, Jessica S.
; APPLICANT: Goddard, Audrey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: INTERLEUKIN-22 POLYPEPTIDES, NUCLEIC ACIDS ENCODING
; FILE REFERENCE: P2806-1 (US)
; CURRENT APPLICATION NUMBER: US/09/870,574
; CURRENT FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: US 60/169,495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 7
; SEQ ID NO 2
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-870-574-2

Query Match      0.7%; Score 8; DB 4; Length 179;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160
DB 21 LLLLLLV 28

RESULT 37
US-09-252-991A-25905
; Sequence 25905, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

```

```

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25905
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25905

Query Match      0.7%; Score 8; DB 4; Length 183;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1171 SARRRRREP 1178
DB 21 SARRRRREP 28

RESULT 38
US-09-040-483-5
; Sequence 5, Application US/09040483
; Patent No. 6143867
; GENERAL INFORMATION:
; APPLICANT: Akerblom, Ingrid E.
; TITLE OF INVENTION: NOVEL HUMAN EOSINOPHIL-
; TITLE OF INVENTION: DERIVED BASIC PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,483
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/740,036
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0133 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 544241
US-09-040-483-5

Query Match      0.7%; Score 8; DB 3; Length 234;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1153 LLLALLV 1160  
|||||  
Db 4 LLLALLV 11

RESULT 39  
US-08-740-036-5  
; Sequence 5, Application US/08740036  
; Patent No. 6492507  
; GENERAL INFORMATION:  
; APPLICANT: Akribiom, Ingrid B.  
; TITLE OF INVENTION: NOVEL HUMAN EOSINOPHIL-  
; DERIVED BASIC PROTEIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/740,036  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0133 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 234 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 544241  
; US-08-740-036-5

Query Match 0.7%; Score 8; DB 4; Length 234;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160  
|||||  
Db 4 LLLALLV 11

RESULT 40  
US-08-118-270-28  
; Sequence 28, Application US/08118270  
; Patent No. 5508384  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Randall B.  
; APPLICANT: Schuster, David I.  
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
; DERIVED BASIC PROTEIN  
; NUMBER OF SEQUENCES: 348  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300

; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/118,270  
; FILING DATE: 09-SEP-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/943,236  
; FILING DATE: 10-SEP-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Townsend, Kevin G.  
; REGISTRATION NUMBER: 34,033  
; REFERENCE/DOCKET NUMBER: MURPHY-2A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; TELEX: 248633  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 315 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-118-270-28

Query Match 0.7%; Score 8; DB 1; Length 315;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLV 1161  
|||||  
Db 48 LLLALLV 55

RESULT 41  
PCT-US93-08528-28  
; Sequence 28, Application PC/TUS9308528  
; GENERAL INFORMATION:  
; APPLICANT: New York University  
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
; DERIVED BASIC PROTEIN  
; NUMBER OF SEQUENCES: 348  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/08528  
; FILING DATE: 09-SEP-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/943,236  
; FILING DATE: 10-SEP-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Townsend, Kevin G.  
; REGISTRATION NUMBER: 34,033  
; REFERENCE/DOCKET NUMBER: MURPHY-2 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197

TELEPAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 315 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US93-08528-28

Query Match 0.7%; Score 8; DB 5; Length 315;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 48 LLLALLVL 55

## RESULT 42

US-09-489-039A-7417  
Sequence 7417, Application US/09489039A  
Patent No. 6610836

## GENERAL INFORMATION:

APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
FILE REFERENCE: 2709-2004001  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 7417

LENGTH: 345  
TYPE: BPT  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-7417

Query Match 0.7%; Score 8; DB 4; Length 345;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLALL 1158  
DB 76 GGLLLALL 83

## RESULT 43

US-09-046-736-4  
Sequence 4, Application US/09046736  
Patent No. 6090582

## GENERAL INFORMATION:

APPLICANT: KIKLY, KRISTINE  
APPLICANT: ERICKSON-MILLER, CONNIE  
TITLE OF INVENTION: Sialoadhesin Family Member-3  
NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ratner & Prestia  
STREET: P.O. Box 980  
CITY: Valley Forge  
STATE: PA  
COUNTRY: USA

ZIP: 19482

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/046,736  
FILING DATE: 24-MAR-1998

CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/041,885  
FILING DATE: 02-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Prestia, Paul F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GH-50019  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0701  
TELEX:

INFORMATION FOR SEQ ID NO: 4:

## SEQUENCE CHARACTERISTICS:

LENGTH: 374 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-046-736-4

Query Match 0.7%; Score 8; DB 3; Length 374;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 379 GVLGAVG 386  
DB 257 GVLGAVG 264

## RESULT 44

US-07-626-618A-17  
Sequence 17, Application US/07626618A  
Patent No. 5422265

## GENERAL INFORMATION:

APPLICANT: Van Tol, Hubert H.M.  
APPLICANT: Civelli, Olivier  
TITLE OF INVENTION: A No. 5422265el Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegratti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA

ZIP: 60606

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/07/626,618A

APPLICATION NUMBER: US/07/626,618A

FILING DATE: 7 DEC 1990

## CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: No. 5422265nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 810-221-8317

INFORMATION FOR SEQ ID NO: 17:

## SEQUENCE CHARACTERISTICS:

LENGTH: 387 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

US-07-626-618A-17

Query Match

0.7%; Score 8; DB 1; Length 387;

Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 45  
US-08-086-439C-3  
Sequence 3, Application US/08086439C  
Patent No. 5468615  
GENERAL INFORMATION:  
APPLICANT: Chio, Christopher L.  
APPLICANT: Huff, Rita M.  
TITLE OF INVENTION: A Synthetic Gene for D4 Dopamine  
RECEPTORS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Upjohn Company, Corp. Intellectual  
ADDRESSEE: Property Law  
STREET: 301 Henrietta Street  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/086.439C  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Darnley Jr., James D.  
REGISTRATION NUMBER: 33,673  
REFERENCE/DOCKET NUMBER: 4700  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 616-385-5210  
TELEFAX: 616-385-6897  
TELEX: 224401  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-086-439C-3

Query Match 0.7%; Score 8; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 46  
US-08-056-051-2  
Sequence 2, Application US/08056051  
Patent No. 5516683  
GENERAL INFORMATION:  
APPLICANT: Grandy, David K  
APPLICANT: Bunzow, James R  
APPLICANT: Civeilli, Olivier  
APPLICANT: Van Tol, Hubert H.-M.  
TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/056.051  
FILING DATE: 19930429  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5516683nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092-C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 910-221-5317  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-056-051-2

Query Match 0.7%; Score 8; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 47  
US-07-928-611-18  
Sequence 18, Application US/07928611  
Patent No. 5569601  
GENERAL INFORMATION:  
APPLICANT: Van Tol, Hubert H.M.  
APPLICANT: Civeilli, Olivier  
TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Allegretti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/928.611  
FILING DATE: 19920810  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5569601nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 810-221-8317

INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-928-611-18

Query Match 0.7%; Score 8; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 48

US-08-333-977-17

Sequence 17, Application US/08333977

Patent No. 5594108

GENERAL INFORMATION:

APPLICANT: Van Tol, Hubert H.M.

APPLICANT: Civeilli, Olivier

TITLE OF INVENTION: A No. 5594108el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/333,977

FILING DATE: 03-NOV-1994

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/626,618

FILING DATE: 7 DEC 1990

ATTORNEY/AGENT INFORMATION:

NAME: No. 5594108nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 810-221-8317

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 387 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-333-977-17

Query Match 0.7%; Score 8; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 49

US-08-449-153-1

Sequence 1, Application US/08449153  
Patent No. 5686255  
GENERAL INFORMATION:  
APPLICANT: Deth, Richard C.  
TITLE OF INVENTION: Compositions and Methods for Diagnosing  
TITLE OF INVENTION: Schizophrenia  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,153  
FILING DATE: 24-MAY-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Cartoll, Alice O.

REGISTRATION NUMBER: 33,542

REFERENCE/DOCKET NUMBER: RCD95-02

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 387 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-449-153-1

Query Match 0.7%; Score 8; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 50

US-08-434-877-3

Sequence 3, Application US/08434877

Patent No. 5721132

GENERAL INFORMATION:

APPLICANT: Chio, Christopher L.

APPLICANT: Huff, Rita M.

TITLE OF INVENTION: A Synthetic Gene for D4 Dopamine

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: The Upjohn Company, Corp. Intellectual

ADDRESS: Property Law

STREET: 301 Henrietta Street

CITY: Kalamazoo

STATE: Michigan

COUNTRY: USA

ZIP: 49001

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette (DS,HD)

COMPUTER: Gateway 2000, P5-90

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/434,877

FILING DATE: 1 July 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Darnley Jr., James D.  
REGISTRATION NUMBER: 33,673  
REFERENCE/DOCKET NUMBER: 4700 DVI  
TELEPHONE: 616-385-5210  
TELEFAX: 616-385-6897  
TELEX: 224401  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-434-877-3

Query Match 0.7%; Score 8; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 51  
US-08-449-651-1  
Sequence 1, Application US/08449651  
Patent No. 5738998  
GENERAL INFORMATION:  
APPLICANT: Deth, Richard C.  
TITLE OF INVENTION: Compositions and Methods for Diagnosing  
TITLE OF INVENTION: Schizophrenia  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,651  
FILING DATE: 24-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Alice O.  
REGISTRATION NUMBER: 33,542  
REFERENCE/DOCKET NUMBER: RCD95-02  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-449-651-1

Query Match 0.7%; Score 8; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 52  
US-08-487-811A-18  
Sequence 18, Application US/08487811A  
Patent No. 5883226  
GENERAL INFORMATION:  
APPLICANT: Civelli, Olivier  
APPLICANT: Van Tol, Hubert H.M.  
TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McDonnell Bohnen Hulbert & Berghoff  
STREET: 300 South Wacker Drive  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,811A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5883226nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092-L  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
TELEX:  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-487-811A-18

Query Match 0.7%; Score 8; DB 2; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 53  
US-08-833-703-1  
Sequence 1, Application US/08833703  
Patent No. 6080549  
GENERAL INFORMATION:  
APPLICANT: DETH, RICHARD  
TITLE OF INVENTION: METHODS AND MATERIALS FOR THE DIAGNOSIS AND  
TITLE OF INVENTION: TREATMENT OF SCHIZOPHRENIA AND RELATED DISORDERS  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
STREET: Ten Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette



```
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,703
; FILING DATE: 08-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Heine, Ph.D., Holliday C
; REGISTRATION NUMBER: 34,346
; REFERENCE/DOCKET NUMBER: NU-431XX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542 2290
; TELEFAX: 617-451 0313
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-833-703-1

Query Match      0.7%; Score 8; DB 3; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 54
US-08-475-742-4
; Sequence 4, Application US/08475742
; Patent No. 6121015
; GENERAL INFORMATION:
; APPLICANT: O'Malley, Karen L
; APPLICANT: Todd, Richard D
; TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4 Receptor
; FILE REFERENCE: WU 102 CON DIV
; CURRENT APPLICATION NUMBER: US/08/475,742
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: US 08/261,293
; EARLIER FILING DATE: 1994-06-16
; EARLIER APPLICATION NUMBER: US 08/014,013
; EARLIER FILING DATE: 1993-01-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(387)
; OTHER INFORMATION: Human D4 Receptor Protein
; US-08-475-742-4
```

```
Query Match      0.7%; Score 8; DB 3; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88
```

```
RESULT 55
US-09-060-694-18
```

```
; Sequence 18, Application US/09060694
; Patent No. 6203998
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 6203998el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,694
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6203998nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-MM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-060-694-18

Query Match      0.7%; Score 8; DB 3; Length 387;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 56
US-09-378-074-18
; Sequence 18, Application US/09378074
; Patent No. 6437114
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/378,074
; FILING DATE: 20-Aug-1999
; CLASSIFICATION: <Unknown>
```

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/928,611  
;; FILING DATE: <Unknown>  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: No. 6437114nan, Kevin E  
;; REGISTRATION NUMBER: 35,303  
;; REFERENCE/DOCKET NUMBER: 90,1092-B  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312-715-1000  
;; TELEFAX: 312-715-1234  
;; TELEX: 810-221-8317  
;; INFORMATION FOR SEQ ID NO: 18:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 387 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-09-378-074-18  
  
Query Match 0.7%; Score 8; DB 4; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88  
  
RESULT 57  
US-08-261-293-4  
; Sequence 4, Application US/08261293  
; Patent No. 5486310  
; GENERAL INFORMATION:  
; APPLICANT: O'Malley, Karen L.  
; APPLICANT: Todd, Richard D.  
; TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4  
; TITLE OF INVENTION: Receptor  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 1100 Peachtree Street, Suite 2800  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: U.S.  
; ZIP: 30309-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/261,293  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/014,013  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: WU 102  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 815-6524  
; TELEFAX: (404) 815-6555  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 387 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO

;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: N-terminal  
;; ORIGINAL SOURCE:  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: Protein  
;; LOCATION: 1..387  
;; OTHER INFORMATION: /note= "Human D4 Receptor Protein"  
US-08-261-293-4  
  
Query Match 0.7%; Score 8; DB 4; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88  
  
RESULT 58  
PCT-US93-07370-18  
; Sequence 18, Application PC/TUS9307370  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses  
; NUMBER OF SEQUENCES: 22  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/07370  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 387 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US93-07370-18  
  
Query Match 0.7%; Score 8; DB 5; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88  
  
RESULT 59  
US-08-056-051-4  
; Sequence 4, Application US/08056051  
; Patent No. 5516683  
; GENERAL INFORMATION:  
; APPLICANT: Grandy, David K  
; APPLICANT: Bunzow, James R  
; APPLICANT: Civelli, Olivier  
; APPLICANT: Van Tol, Hubert H.-M.  
; TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and Uses  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Allegretti & Witcoff, Ltd.  
; STREET: 10 South Wacker Drive, Suite 3000  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/056,051  
;; FILING DATE: 19930429  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: NO. 5516683nan, Kevin E  
;; REGISTRATION NUMBER: 35,303  
;; REFERENCE/DOCKET NUMBER: 90,1092-C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312-715-1000  
;; TELEFAX: 312-715-1234  
;; TELEX: 910-221-5317  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 419 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-056-051-4

Query Match 0.7%; Score 8; DB 1; Length 419;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1154 LLLALLVL 1161  
|||  
DB 81 LLLALLVL 88

RESULT 60  
US-07-928-611-20  
;; Sequence 20, Application US/07928611  
;; Patent No. 559601  
;; GENERAL INFORMATION:  
;; APPLICANT: Van Tol, Hubert H.M.  
;; TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses  
;; NUMBER OF SEQUENCES: 22  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Allegretti & Witcoff, Ltd.  
;; STREET: 10 South Wacker Drive, Suite 3000  
;; CITY: Chicago  
;; STATE: Illinois  
;; COUNTRY: USA  
;; ZIP: 60606  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/928,611  
;; FILING DATE: 19920810  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: NO. 5569601nan, Kevin E  
;; REGISTRATION NUMBER: 35,303  
;; REFERENCE/DOCKET NUMBER: 90,1092-B  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312-715-1000  
;; TELEFAX: 312-715-1234  
;; TELEX: 810-221-8317  
;; INFORMATION FOR SEQ ID NO: 20:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 419 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-07-928-611-20

Query Match 0.7%; Score 8; DB 1; Length 419;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1154 LLLALLVL 1161  
|||  
DB 81 LLLALLVL 88

RESULT 61  
US-08-487-811A-20  
;; Sequence 20, Application US/08487811A  
;; Patent No. 5883226  
;; GENERAL INFORMATION:  
;; APPLICANT: Civelli, Olivier  
;; APPLICANT: Van Tol, Hubert H.M.  
;; TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses  
;; NUMBER OF SEQUENCES: 24  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff  
;; STREET: 300 South Wacker Drive  
;; CITY: Chicago  
;; STATE: IL  
;; COUNTRY: USA  
;; ZIP: 60606  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/487,811A  
;; FILING DATE: 07-JUN-1995  
;; CLASSIFICATION: 530  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: No. 5883226nan, Kevin E  
;; REGISTRATION NUMBER: 35,303  
;; REFERENCE/DOCKET NUMBER: 90,1092-L  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312-913-0001  
;; TELEFAX: 312-913-0002  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 20:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 419 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-487-811A-20

Query Match 0.7%; Score 8; DB 2; Length 419;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1154 LLLALLVL 1161  
|||  
DB 81 LLLALLVL 88

RESULT 62  
US-09-060-694-20  
;; Sequence 20, Application US/09060694  
;; Patent No. 620398  
;; GENERAL INFORMATION:  
;; APPLICANT: Civelli, Olivier  
;; APPLICANT: Van Tol, Hubert H.M.  
;; TITLE OF INVENTION: A No. 620398el Human Dopamine Receptor and Uses  
;; NUMBER OF SEQUENCES: 24  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff  
;; STREET: 300 South Wacker Drive  
;; CITY: Chicago  
;; STATE: IL  
;; COUNTRY: USA  
;; ZIP: 60606  
;; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/060,694  
FILING DATE: 15-APR-1998  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6203998nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092-MM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
TELEX:

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:  
LENGTH: 419 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-060-694-20

Query Match 0.7%; Score 8; DB 3; Length 419;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161  
|||||  
Db 81 LLLALLVL 88

## RESULT 63

US-09-378-074-20  
Sequence 20, Application US/09378074  
Patent No. 6437114  
GENERAL INFORMATION:

APPLICANT: Van Tol, Hubert H.M.

Civelli, Olivier

TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/378,074

FILING DATE: 20-AUG-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/928,611

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: No. 6437114nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-B

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 810-221-8317

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:  
LENGTH: 419 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-378-074-20

Query Match 0.7%; Score 8; DB 4; Length 419;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161  
|||||  
Db 81 LLLALLVL 88

## RESULT 64

PCT-US93-07370-20  
Sequence 20, Application PC/TUS9307370  
GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 22

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/07370

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 419 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US93-07370-20

Query Match 0.7%; Score 8; DB 5; Length 419;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161  
|||||  
Db 81 LLLALLVL 88

## RESULT 65

US-09-252-991A-17394  
Sequence 17394, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONA

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 17394

LENGTH: 453

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-17394

Query Match 0.7%; Score 8; DB 4; Length 453;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1150 LGGLLLLA 1157  
|||||

Db 426 LGGLLLLL 433

RESULT 66

US-08-056-051-6

Sequence 6, Application US/08056051

Patent No. 5516683

GENERAL INFORMATION:

APPLICANT: Grandy, David K

APPLICANT: Bunzow, James R

APPLICANT: Civeilli, Olivier

APPLICANT: Van Tol, Hubert H.-M.

TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/056,051

FILING DATE: 19930429

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 5516683nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-C

TELEPHONE: 312-715-1234

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-056-051-6

Query Match 0.7%; Score 8; DB 1; Length 467;

Best Local Similarity 100.0%; Pred. No. 57;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1154 LLLALLVL 1161

Db 81 LLLALLVL 88

RESULT 68

US-08-487-811A-22

Sequence 22, Application US/08487811A

Patent No. 5883228

GENERAL INFORMATION:

APPLICANT: Civeilli, Olivier

APPLICANT: Van Tol, Hubert H.M.

TITLE OF INVENTION: A No. 5883228el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff

STREET: 300 South Wacker Drive

CITY: Chicago

STATE: IL

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,811A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: No. 5883228nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-L

TELEPHONE: 312-913-0001

TELEFAX: 312-913-0002

TELEX:

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-487-811A-22

Db 426 LGGLLLLL 433

RESULT 67

US-07-928-611-22

Sequence 22, Application US/07928611

Patent No. 5569601

GENERAL INFORMATION:

APPLICANT: Van Tol, Hubert H.M.

APPLICANT: Civeilli, Olivier

TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/928,611

FILING DATE: 19920810

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 5569601nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-B

TELEPHONE: 312-715-1234

TELEFAX: 312-715-1234

TELEX: 810-221-8317

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-928-611-22

Query Match 0.7%; Score 8; DB 1; Length 467;

Best Local Similarity 100.0%; Pred. No. 57;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1154 LLLALLVL 1161

Db 81 LLLALLVL 88

RESULT 67

US-07-928-611-22

Sequence 22, Application US/07928611

Patent No. 5569601

GENERAL INFORMATION:

APPLICANT: Van Tol, Hubert H.M.

APPLICANT: Civeilli, Olivier

TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/928,611

FILING DATE: 19920810

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 5569601nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-B

TELEPHONE: 312-715-1234

TELEFAX: 312-715-1234

TELEX: 810-221-8317

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-928-611-22

```

Query Match      0.7%; Score 8; DB 2; Length 467;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1154 LLLALLVL 1161
Db      81 LLLALLVL 88

RESULT 69
US-09-046-736-2
; Sequence 2, Application US/09046736
; Patent No. 6090582
; GENERAL INFORMATION:
; APPLICANT: KIKLY, KRISTINE
; APPLICANT: ERICKSON-MILLER, CONNIE
; TITLE OF INVENTION: Sialoadhesin Family Member-3
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P. O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,736
; FILING DATE: 24-MAR-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/041,885
; FILING DATE: 02-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-50019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-046-736-2

Query Match      0.7%; Score 8; DB 3; Length 467;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      379 GVLGAVG 386
Db      350 GVLGAVG 357

RESULT 70
US-09-060-694-22
; Sequence 22, Application US/09060694
; Patent No. 6203998
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,694
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6203998nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-MM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-060-694-22

Query Match      0.7%; Score 8; DB 3; Length 467;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1154 LLLALLVL 1161
Db      81 LLLALLVL 88

RESULT 71
US-09-378-074-22
; Sequence 22, Application US/09378074
; Patent No. 6437114
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/378,074
; FILING DATE: 20-AUG-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/928,611
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6437114nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-B
; TELECOMMUNICATION INFORMATION:

```

TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 810-221-8317  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-378-074-22

Query Match 0.7%; Score 8; DB 4; Length 467;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 72  
PCT-US93-07370-22  
Sequence 22, Application PC/TUS9307370  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/07370  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US93-07370-22

Query Match 0.7%; Score 8; DB 5; Length 467;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 73  
US-08-485-618-103  
Sequence 103, Application US/08485618  
Patent No. 5728533  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5728533rel Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

Query Match 0.7%; Score 8; DB 1; Length 494;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLLLL 1158  
DB 442 GGLLLLLL 449

RESULT 74  
US-08-605-672-103  
Sequence 103, Application US/08605672  
Patent No. 5817515  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/605,672  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:

TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 810-221-8317  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-378-074-22

Query Match 0.7%; Score 8; DB 4; Length 467;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 72  
PCT-US93-07370-22  
Sequence 22, Application PC/TUS9307370  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/07370  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US93-07370-22

Query Match 0.7%; Score 8; DB 5; Length 467;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 73  
US-08-485-618-103  
Sequence 103, Application US/08485618  
Patent No. 5728533  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5728533rel Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

Query Match 0.7%; Score 8; DB 1; Length 494;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLLLL 1158  
DB 442 GGLLLLLL 449

RESULT 74  
US-08-605-672-103  
Sequence 103, Application US/08605672  
Patent No. 5817515  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/605,672  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:

NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32684  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 494 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-605-672-103

Query Match 0.7%; Score 8; DB 2; Length 494;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLAL 1158  
DB 442 GGLLLAL 449

## RESULT 75

US-08-482-293A-103  
Sequence 103, Application US/08482293A  
Patent No. 5831029  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5831029e1 Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,293A  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32684  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 494 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-482-293A-103

Query Match 0.7%; Score 8; DB 2; Length 494;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLAL 1158  
DB 442 GGLLLAL 449

Search completed: June 24, 2004, 18:06:46  
Job time : 25 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 17:53:06 ; Search time 27 Seconds

(without alignments)

4232.427 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 6224

Sequence: 1 MDLPRGLVVALSLWPQFT.....FRSARRRREPGLDTPKYLE 1188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 9619526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID    | Description        |
|------------|--------|-------------|--------|----------|--------------------|
| 1          | 2159   | 34.7        | 1151   | 2 A45226 | integrin alpha-1 c |
| 2          | 2149.5 | 34.5        | 1180   | 2 A35854 | integrin alpha-1 c |
| 3          | 1910.5 | 30.7        | 1178   | 2 A41442 | VLA-2 protein homo |
| 4          | 1901.5 | 30.6        | 1170   | 2 I45914 | integrin alpha 2 s |
| 5          | 1863   | 29.9        | 1181   | 2 A33998 | integrin alpha-2 c |
| 6          | 1146.5 | 18.4        | 1170   | 2 S03308 | cell surface glyco |
| 7          | 1135.5 | 18.2        | 1163   | 1 RWHU1C | cell surface glyco |
| 8          | 1104   | 17.7        | 1153   | 2 S00551 | leukocyte surface  |
| 9          | 1094   | 17.6        | 1153   | 1 RWHU1B | cell surface glyco |
| 10         | 1073.5 | 17.2        | 1163   | 2 I54126 | lymphocyte fuction |
| 11         | 924.5  | 14.9        | 1179   | 2 A53213 | integrin alpha-E c |
| 12         | 799.5  | 12.8        | 1035   | 2 I58409 | integrin alpha-9 c |
| 13         | 789    | 12.7        | 272    | 2 A55348 | integrin alpha-1 c |
| 14         | 738    | 11.9        | 1041   | 2 T31437 | integrin alpha-1 c |
| 15         | 719.5  | 11.6        | 1054   | 2 JCT294 | alphan integrin -  |
| 16         | 693.5  | 11.1        | 1039   | 2 A41131 | lymphocyte-Peyer's |
| 17         | 672.5  | 10.8        | 1038   | 2 S06046 | integrin alpha-4 c |
| 18         | 607.5  | 9.8         | 1034   | 2 A36108 | integrin alpha-V c |
| 19         | 606    | 9.7         | 1049   | 2 A27079 | fibronectin recept |
| 20         | 605    | 9.7         | 1091   | 2 A41543 | integrin alpha-6 c |
| 21         | 604.5  | 9.7         | 1053   | 2 S44250 | integrin alpha-5 c |
| 22         | 604    | 9.7         | 1051   | 2 A40021 | integrin alpha-3 c |
| 23         | 602.5  | 9.7         | 1073   | 2 B36429 | integrin alpha-6 c |
| 24         | 594    | 9.5         | 1053   | 2 I55534 | VLA-3 alpha subuni |
| 25         | 587    | 9.4         | 1051   | 2 A35761 | cell surface glyco |
| 26         | 584.5  | 9.4         | 1044   | 2 T10050 | integrin alpha-v c |
| 27         | 568    | 9.1         | 1072   | 2 A38457 | integrin alpha-6 c |
| 28         | 566.5  | 9.1         | 1146   | 2 S40311 | integrin - fruit f |
| 29         | 563    | 9.0         | 1044   | 2 S16516 | integrin alpha-8 c |

546.5 8.8 1226 2 S44834  
544 8.7 1115 2 T09403  
539 8.7 1115 2 T09433  
537 8.6 1048 2 A27421  
536 8.6 1045 2 S60571  
530.5 8.5 1137 2 UC5950  
523.5 8.4 1135 2 I61186  
515.5 8.3 1139 2 S28277  
510 8.2 1037 2 A60163  
502.5 8.1 1106 2 S38783  
478.5 7.7 1039 2 A34269  
470 7.6 1394 2 A29637  
429 6.9 1086 2 T18523  
420 6.7 191 2 I47230  
349.5 5.6 764 2 I36916  
260.5 4.2 3124 2 A40020

## ALIGNMENTS

### RESULT 1

A45226

integrin alpha-1 chain - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 15-Sep-2003

C:Accession: A45226

R:Brisesewitz, R.; Epstein, M.R.; Marcantonio, E.B.

J. Biol. Chem. 268, 2989-2996, 1993

A:Title: Expression of native and truncated forms of the human integrin alpha 1 subun

A:Reference number: A45226; MUID:93155124; PMID:8428973

A:Accession: A45226

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1151 <BR1>

A:Experimental source: hepatoblastoma cell line HepG2

A>Note: sequence extracted from NCBI backbone (NCBIP:124326)

P:142-317/Domain: von Willebrand factor type A repeat homology <YMAL>

Query Match 34.7%; Score 2159; DB 2; Length 1151;

Best Local Similarity 38.3%; Pred. No. 8e-136;

Matches 457; Conservative 235; Mismatches 417; Indels 84; Gaps 20;

|    |     |  |     |
|----|-----|--|-----|
| QY | 23  | FNMUTRPRVTPGSRTPFFGTVOQHDISGNKWLAVVGLAPLETNGYQKTGDVYKCPVHGN  | 82  |
| DB | 1   | FNDVKNMTSGPVEDMFEGTVQYENEGKWLIGSLVQPKRKTGDVYKCPVGRGE         | 60  |
| QY | 83  | ---CTKMLG-RVTLNVSERKDNRLGLSLATPKDNPGLACPLWSHEGSSVYTRGM       | 138 |
| DB | 61  | SLPCVKLDLPVNTSIPNVTYKENTFGSTLTATNP-NGGPLACGLYAYRCOHLHYTTGI   | 119 |
| QY | 139 | CSRYNSNFRFKTAPALQRCQTYMDIVIVLDGNSIYPMWVEVQVHFLINI LKKFYIGPQQ | 198 |
| DB | 120 | CSDVSPTFQVNSIAP-VQECSTQLDIVIVLDGNSIYPMWSDVTAFLNDLLKEMDIGPKQ  | 178 |
| QY | 199 | IQGVGVQGVGDVHFFHNDYRSVKD VVEAAASHIEQGGTTRTAFTAGIEFARSEAP--QK | 256 |
| DB | 179 | TQGVGVQGVGVNTHEFNLYKTSSTEEVLAARKIVQRGGRQMTALGDTARKEAPTAR     | 238 |
| QY | 257 | GGRGAKKVMIVITDGHSDSPDLKVKIQQSERDNTVRYAVAVLGYNNRRGINPTFLNE    | 316 |
| DB | 239 | GARGVAKVMIVITDGHSDHNRKLVQDCEDENIQPFTAILGSLYNNRNLSTEFVBE      | 298 |
| QY | 317 | IKYIASDPDDKHPNVTDEAALKDIVDALGRIFSELT-KNNETSFGLEMSQTGFSHV     | 375 |
| DB | 299 | IKSIASEPTEKHFFNVSDALAVITVKTLGERIFALEATADQSAASFEMMSQTGFSAHY   | 358 |
| QY | 376 | VEDGVLLGAVGAYDMGVALKETSAGKVIPLRESVYLKEFFELKNHGYLVGYTVSVVSS   | 435 |
| DB | 359 | SDQWMLGAVGAYDMGVTVMQKASQIIIPNTTFNVSTKKNELASLYLGYTVNSATAS     | 418 |
| QY | 436 | ROGRVYVAGAPRNFHTGKVLFTMHNNRSLTIHQAMRQQQIGSYFGSITSVIDDGDVYT   | 495 |

Db 419 SGDVLXIAGQPRYHNTQVILIRYMDG-NIKILQTLGSGEIGSYFGSILTTTDDKDSNT 477  
Qy 496 DVLIVGAPMYF-NEGHERGKVVYVELAQRNFVNGTLK-----DSHSYQN--- 539  
Db 478 DILLVGAPMYNGTEKEBQGVYVALNQTFRYQMSLEPIKOTCCSSRQHSNCTTINKNE 537  
Qy 540 ---ARFGSSIASVRDLNODSYNDVVVCGAPLEDNHAGAIYIFHGFGRSILKTPKORTAGE 596  
Db 538 PCGARGTAAVAKDLNODFNIVIGAPLEDHGGAVIYHSGKTIKRYAQRIPSGG 597  
Qy 597 LATGLQYFGSGIHGQDLNEDGLIDLAVGALNAVILWSPVQVQINASLHPEPSKINIFH 656  
Db 598 DGKTLKFFGSGIHGEMDNGDGLTDTVTIGLGAALFWSRDVAVVYVWTFNPEPNKNIQK 657  
Qy 657 RDCRSRGRATCLAAFLCFTPIFLAPHFTQTTVGIRYNATMDERRVTPRAHLDGGDRFT 716  
Db 658 KNCHEGKEVTCINAVCFVKLSKEDTIYADLOVRLDLSRQISRSFFSGTGERKV 717  
Qy 717 NRAVLSSGQELCERINFHVLDTADYVYKPTVTFVSEYSLDDPKGPMDDGWPTTLKXVP 776  
Db 718 QRNITVRKSB--CTGSGFVMLDKHDPODSVRITLDFNLTDPNPENGVLDDSLPNSVHEYIP 775  
Qy 777 FVNGCRDEHCVPLVLDARSDDLPTAMEYQORVLRKPAQDCSAYTLSPDTTFFLIESTRQ 836  
Db 776 FAKDCGNKEKICISDL-----SUHVATTEKDLILVRSND 809  
Qy 837 RVAVEATLNRGENAYSTVLNISQSANLQFASL--IQKEDSDGSIBCVWEERELQKOVCN 894  
Db 810 KENVSLVTKTKDSAYNTRTIVHSPNLVPSGIEAIQKSCSN-----HNITCK 859  
Qy 895 VSYPPFFAKAVAFRLDSRESKSIPLHLEIEIAAGSDSNERSTKEDNVAPLPHLYE 954  
Db 860 VGYPFLRGWVTKILFQNTSVLMENVTIYISATSDSEEPETLSDNVNISIPVKYE 919  
Qy 955 ADVLFTRSSLSHVEVKLNS---LERYDIGPFPFSCIFRIQNLGLFPIHGMKMITIP 1010  
Db 920 VGLQFVSASBYHISIAANETVPEVINSTEDIGNEINIFYLIRKSGSPPELKLISFP 979  
Qy 1011 IATRSGRNLLKRLDFTDEANTSC-----NIGNSTEYPTPTVEEDLRAPOLNSNSD 1064  
Db 980 NMTSNGYFVL-YPTGUSSENANCRPHIFEDPDSINSKGVKTTSTDLKRGTLDCNTCK 1038  
Qy 1065 VWSINCIRLVPMQIEINFHLLGNLWLSKALKYKSKIMVNAALORQHSPIFREEDP 1124  
Db 1039 FATITCNLTSSDISQVNSLI--LWKPTFIKSYFSSNLIRGEL-RSENASLVSSNQ 1095  
Qy 1125 SRQIEFISKQE-DWQVPIMIVGSTGGLLLALLVLRKLGKGFPSARRR 1176  
Db 1096 KRELAIQISKGLPGRVPLWILLAPAGLLMLLILALWIKIGFRRPLKXK 1148

RESULT 2  
A35854  
Integrin alpha-1 chain precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 23-Oct-1990 #sequence\_revision 13-Sep-1991 #text\_change 15-Sep-2003  
C:Accession: A35854; S11243  
R:Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J. Cell Biol. 111, 705-720, 1990  
A:Title: Molecular cloning of the rat integrin alpha-1 subunit: a receptor for laminin a  
A:Reference number: A35854; MUID:90338125; PMID:2380249  
A:Accession: A35854  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1180 <ICN>  
A:Cross-references: GB:X52140; NID:G56493; PIDN:C2A36384.1; PID:G56494  
C:Keywords: cell adhesion; cytoskeleton; transmembrane protein  
F:170-345/Domain: von Willebrand factor type A repeat homology <WA2>

Query Match 34.5%; Score 2149.5; DB 2; Length 1180;  
Best Local Similarity 38.0%; Pred. No. 3.6e-135;  
Matches 463; Conservative 242; Mismatches 405; Indels 107; Gaps 24;

Qy 11 WALSLWPGPTDTFNMDTKPRVIPCGRSTAPFGYTVQOHDISGNKWLVLVVCAPLETNGYOKT 70  
Db 17 WLLTVILGFCVFNVDVFKSMSFSGPVEDMFQYTVQOYENEBGKXVLIGSLVWGPKART 76  
Qy 71 GDVYKCPVTHGN---CTKLNGL-RVTLNVSERKDNMLGLSLATNPKNDSFLACPLMS 126  
Db 77 GDVYKCPVGRBRAMPCKVLDLPLVNTSIPNVTIENKMTFGSLVATNP--NGGFLACGELYA 135  
Qy 127 HEGSSSYTTGMCSSRVNSNFRFSKTVAPALORCQYMDIVIVLDGNSNIYPMVEVQHFLI 186  
Db 136 YKCGHLRTTIGCSUSVPTFPQVNSFAP-VQECSTQLDVIIVLDGNSNIYPMVESVIAFLN 194  
Qy 187 NILKFPYIPGQIQGVVQYGEDVVHFLANDYRSVKDVVEAASHIEORGGTETRTAFIGI 246  
Db 195 DLLKMDIGPKTQGVIGVYGENVTHFNLANKYSSTBEVLVAANKIGRQGGIQTMTALGI 254  
Qy 247 EPARSEAP---QKGRGKAKKVMIVITDGHSDSDPLEKVIQOSERDNVTRVAVAVLGYN 304  
Db 255 DTARKEATFEARGARGVKVWVITDGHSDHNYRLKQVQDCEDENQRFSAIALGHFN 314  
Qy 305 REGINPFLNEIKYIASDPDDPKHPFNVTDSAAKDIYDALGDRIPLSLEGT-NKNETSFG 363  
Db 315 RGNLSTEFVBEIKSIASEPTEKHFNVSDLELALVTIVKALGERIFALEATADQSAASPE 374  
Qy 364 LMSOTGSSSHVVVGVLLGAVDNGAVLKETSAGKVIPLRESYLKEPPEELKNHGA 423  
Db 375 MEMSOTGSAHSQDWMLGAVDNGVVMQKANKQVIPHNTTPTPEAKOMEPPLAS 434  
Qy 424 YLGYTWTISVSSROGRV-YVAGAPFNHTGKVLFTMNNRSLTIHQAMRGQOIGSYFGS 482  
Db 435 YLGYTWS--ATIPGDVLYIAGQRYNHTGVVYKMDG-NINILQTLGGGQIGSYFGS 491  
Qy 483 EITSVIDDGDGVTDLVVGAPMYF-NEGRERKQVYVELRQNRFTVN-----GT 530  
Db 492 VLTITDIDKSDYTDLLVGAPMYGTEKEEQGVVYVAVNQTFRFYQMSLEPIRQCCSS 551  
Qy 531 LKDSHSYQ-----NARFGSSIASVRDLNODSYNDVVVCGAPLEDNHAGAIYIFHGFGRS 583  
Db 552 LKDSCTKENKPEPCGAPFGTAIAVKDLNODFNIVIGAPLEDHGGAVIYHSGKT 611  
Qy 584 ILKTPKORTASBELATGLQYFGCSIHGQDLNEDGLIDLAVGALNAVILWSPVQVQINA 643  
Db 612 IREAVAQRIPTGGDKTKLFFQOSIHGEMDNGDGLTDTVTIGLGAALFWARDVAVVKV 671  
Qy 644 SAHFSPSKINIFHRCKSRGRDATCLAAPLCFTPIFLAPHFTQTTVGIRYNATMDERRYT 703  
Db 672 TWFBNKVNIOKKNCRVGKETVCINATMCFHVKLSKEDSIYEADLOVRLDLSRQI 731  
Qy 704 PRAHLDGGDRFTNRAVLSSQBELCERINFHVLDTADYVYKPTVTFVSEYSLDDPKGPM 763  
Db 732 SRSPFSGTQERKIQRNITVRESE--CIRHSFVMLDKHDPODSVRVRLDNLTDPNPENGVL 789  
Qy 764 DDGWPTTLVSVFPFNGCNEDEHCVPLVLDARSDDLPTAMEYQORVLRKPAQDCSAYTL 823  
Db 790 DDALPNSVHEHIFPAKDCGNKRCISDLTLNVSTT----- 824  
Qy 824 FDTVPIIESTRQVAVATLNRGENAYSTVLNISQSANLQFASL--IQKEDSDGSIB 881  
Db 825 -EKSLLVKSQHKENVSILTVNKGDSAYNTRTVQHSNPLIFSGIEEIKQKSCSN--- 880  
Qy 882 VNEERLQKQCNVSPFPRAKAVAFRLDSFPSKSIPLHLEIEIAAGSDSNERSTKE 941  
Db 881 -----QNITCRVGYPFLRAGETVTFKLIIPQNTSHLSENAIHLISATSDSEEPLES 933  
Qy 942 DNVAFLRFLKYEADVLFRSSSLSHVEVKLNSLERY-----DGIGPPSPSCIFRIQNLGL 997  
Db 934 DNEVNISIPVKYEVGLQFYSSASEHHISVAANETIPEFINSTEDIGNEINVTIRKRGH 993  
Qy 998 PFHGMKMITIPIATRSGRNLLKRLDFTDEANTSCNIGWSTERYPTFVSE--DLRA 1055  
Db 994 FPMPELQLSIFPNLTADGPVLYPIGWSSSD--NVNC-----RPSLDDPFGINGSG 1043  
Qy 1056 POLNHSNSDV-----SINCIRLVPMQIEINFHLLGNLWLSKALKYKS 1100

Db 1044 KWTLSKSEVLKRGITQDCSSCTGCVATITCSLLPSDLQVNVSL--LWKPTFRHPS 1101  
Qy 1101 MKIMVNAALORQFHSFFIFREDPSQIEFEISKQE-DKQVPIIMIVGSGTIGLILALL 1159  
Db 1102 LMLTLRGLKSE-NSSLTSSNRKRELAIQISKGLGPRVPLMWILLISAFAGLLMLL 1160  
Qy 1160 VLALKLGLGFRSARRR 1176  
Db 1161 ILALWKIGFRRPLKK 1177

RESULT 3  
S44142  
VLA-2 protein homolog - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 15-Sep-2003  
C:Accession: S44142  
R:Edelma, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Danjanovich, L.; Latze  
submitted to the EMBL Data Library, January 1994  
A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not  
A:Reference number: S44142  
A:Accession: S44142  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1178 <EDB>  
A:CROSS-references: EMBL:Z29987; NID:G473098; PIDN:CAA92877.1; PID:G473099  
F:169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 30.7%; Score 1910.5; DB 2; Length 1178;  
Best Local Similarity 36.2%; Pred. No. 3.5e-119; Indels 103; Gaps 31;  
Matches 442; Conservative 238; Mismatches 437;

Qy 6 GLVVAWALSMPGETD--TFNMDTRKPRVPGSRPTAFPGTVQOHDISGNKWLWVGAPLE 63  
Db 8 GALLQLQLMLVQGLNCLAYNVGLPGAKIFSPSESGFQYSVQQLTNPGQNMLLVGSPTS 67  
Qy 64 TNGYKTKGVYKCPV--IHGNCITKLN--GRVTLNVSVRKONKRLGLSLATNPKNSPLA 120  
Db 68 GPENRMGDVYKCPVDLPATCEKLNQNSISNWTIKTNMSGLTLTRNPGGGFLT 127  
Qy 121 CSPLWSHEGSSYYTGMCSRNSNFRSKTVAPALQRCQTYMDIVIVLDGNSNTPWVE 180  
Db 128 CGPLWAHQNGYATGICSDVSPDFLTSFPAVQACPSLVDVVVWCDNSNTPWEA 187  
Qy 181 VQFLINILKFIYGPQIQGVQGVGDVWVHEPHLNDYRSVKDVVEAASHTEORGGTET 240  
Db 188 VKNFLVKFTGLDIPKTKQVALIQVANEPIIFNLNDFETKEDWVQATSETROHGGDLT 247  
Qy 241 RTAFGEFARSAFOK--GGRGAKKQVIMIVITDGHSDSPDLKVIQOESRDNVTRYAVA 298  
Db 248 NTFRAIEFADYAYSQTSGRPGATKVMVVTDSHDSGLKTVLQCNDEILRFGIA 307  
Qy 299 VLGYNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAKDIVDALGRIPSLSGTKN 358  
Db 308 VLGYLNRNALDKNLKEIKAIASPTERYFFNVADERALLKAGTAGEQIFSIETVQG 367  
Qy 359 ETSPLGEMSGTGFSSHV--EDGVLLGAVGDMGAVLKETSAGKVIPLRESYLKBPPE 416  
Db 368 GDNFQEMQAQVFSADYAFQNDILMIGAVGADMSGLVQETSHKPV-----PFK 418  
Qy 417 EL-----KNFGAIVGYTTSVSSRQGRVYVAGAPRENHTGKVLFTMNNRSLTIHQ 469  
Db 419 QAQDQVLQDRNHSPLGYSV-RAISTEDGVHFGVAGAPRANYTQIVLVSNVQGNVTIQ 477  
Qy 470 AMRQOIGSYFSGSEITSDIDGVDVLLVAGPMYFNE-GRERGKVVYVELRQNFVYN 528  
Db 478 SHRGDQIGSYPGSVLCSDVDKDTITDVLVGLAPTMYNDLKKBEKGVILFTTIKILNQH 537  
Qy 529 GTLKDSHVSQNAFFSGSIASVRDLNQDSYNDVVVWVGLAPLNDHAGAIYIFPHGFRGSLKTP 588  
Db 538 QFLEGPECTGNARFGSAIALSDINNDGNDVIGSPVENSGAVIYNHGQITRTKY 597

Qy 589 KQKITASELA--TGLQVFGCSIHQDLNEDGLDLAVGALGNNAVILWSRPVVQINASLH 646  
Db 598 SOKILSGNAGFRRLHQLQFGRSLDGYDLNGDSITDVSIGALQVQLWSQSIADVAIEAL 657  
Qy 647 FBSKINIHRDCKRSRDATCAALFLCFTPIFLAPHFOITTVGIRVATWD----BRRY 702  
Db 658 FTPDKITILLNDAK-----ITLKLCTFRABF-RPAGQNNQVAILFNWILDAOCHSRV 708  
Qy 703 TPRAHLDGEGDRFTNRVALLSSGQELACERINFHVLDTADYVYKPTVFSVYSLEDDHGM 762  
Db 709 TSGVTPRENSERFLQKMMVNEVQK-CSEHHSIQKPSDVVNPDLAVDISLENGTSPA 767  
Qy 763 LDDGWPTTLRV-SVFPWNGCNEDEHCVPLDLVLDARSDLPTAMEYCORVLRKPAQCCSAYT 821  
Db 768 L-EAYSETVRFVPSIPFYKCGSGDGCISDILVDQ--QLP----- 804  
Qy 822 LSPDTTVPITESTRQVAVAEATLENGENAYSTVLNISQSANLOFASL1QKEDSDGSIEC 881  
Db 805 -AIQTSFIVSNQKRLTPSVILLKNRGSAYNTVLAEFSENLFASFSPMPVD--GTEV 860  
Qy 882 VNEERRLOKQV-CNVSPFPRAKAVAPRLDSEFSKIFLHLHLELAAGSDNSRERSTK 940  
Db 861 TCEVSSQKSVTCVGYFALKSEQOVTTFINDFNLQNLQAAINFQAFSESOE--TNK 918  
Qy 941 EDNVAPLRFHLKYEADVLTFRSSSLSHVEKLN---SSLERYDGIQPPFCIFRIQ-NL 995  
Db 919 ADNSVSTITPLLYDAEHLTRSTNINFEYISSDENAPSVIKSVEDIGPKF--IFSLKYTA 976  
Qy 996 GLFPIHGMMKTIPIATRSNRLKLRLDFLTDBA--NTSCNWNSTNRYRTPV----- 1048  
Db 977 GSAPVSMALVTIHIPOYTKENPLLYLTGIGTDOAGDISC-----TABINPLKLPHTAPS 1031  
Qy 1049 ----EEDLRAPQLNHSNDVVSINCNIR-LVNPQENFHLGNLRLSLKALYKSKMI 1103  
Db 1032 VSPKNENFRHTKEDUCCRTTSCSNITCWLKOLHMAEYFINVTRVWNTPAASFTQ--V 1089  
Qy 1104 MVNAALQRFHSPIFREEDPSRQIRFEISK-QEDWQVPIIIVGSLTGILLALLALVLA 1162  
Db 1090 QLTAABEIDTNPOLFVIEENAVTIPLMIMKPTKEAEVPTGVIIGSIAGILLAMTAG 1149  
Qy 1163 LRKLGPFRSARRRREPGLDP 1182  
Db 1150 LMKLGFTK--ROYKMGQNP 1167

RESULT 4  
I45914  
Integrin alpha 2 subunit - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 15-Sep-2003  
C:Accession: I45914  
R:Kanata, T.; Puzon, W.; Takada, Y.  
J.Biol. Chem. 269, 9659-9663, 1994  
A:Title: Identification of putative ligand binding sites within the I-domain of integ  
A:Reference number: A54402; MUID:94193647; PMID:7511592  
A:Accession: I45914  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1170 <KAM>  
A:CROSS-references: GB:I25886; NID:G439695; PIDN:AAB59255.1; PID:G439696  
F:161-336/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 30.6%; Score 1901.5; DB 2; Length 1170;  
Best Local Similarity 36.3%; Pred. No. 1.4e-118;  
Matches 437; Conservative 238; Mismatches 437; Indels 91; Gaps 30;

Qy 23 FNMOTRKRPRVPGSRPTAFPGTVQOHDISGNKWLWVGAPLENTGYQKTDGVYKCPV--IH 80  
Db 19 YNVLGPKAKIFSGPSSEQFVAVQOQFNPKGNWLLVGSFPMGPPKRNMGDVKCPVDLST 78  
Qy 81 GNCTKLNIGRVT--LSNVSRKDNRLGLSLATNPKNSFLACPSLWSHEGSSYYTTGMC 139  
Db 79 TTCCKLNLTQSTSMNSNTEMTKNTMSGLTLTRNVGTGGFLTCGPPLWAQCCSQYYTTGVC 138





Db 335 EGTSKODLTSFNNELSSGSLADLSRSHAVGVAGAKDW-----AGGFLLDLKADQL 385  
 Qy 408 -ESYLKEFPBELKHGAYLYGYTTSVVVSQGRVYVAGAPRNFHTGKVLFTM-----HN 461  
 Db 386 DDTFIGNELPTEVRAGLYGYTTLPLPSRQKTSLLASGAPRYQHWGRVLLFPBQPGGCHW 445  
 Qy 462 NRSLLTHQMRGQIGSYGSEITSVGIDCGVTDVLLGAPMYFNEGRGKVVYELR 521  
 Db 446 SQVTH-----GQIGSYGSEIGCGVDVDDGTELLLGAPLYGQR--GRKVFYQRR 500  
 Qy 522 QNFVYNGTLKDSHYSQNARFGSSIASVRDLNDSYNDVVVGVGAPLRDNHAGAIYIFPHFR 581  
 Db 501 QLGFEVSEVLQGDGPGYPLGRFGAITALTDINGDLVDVAVGAPLEEQ--GAVYIFNGRH 558  
 Qy 582 GSILKTPKORITASELATGLOYGCSITHGDLNEDGLDLAVAGLGNNAVILMSRPVIOI 641  
 Db 559 GELSPQSQIRIGTVQLSGIQMFGRSHIGVKDLEGDLADVAVGASQIMVLSRRVDDM 618  
 Qy 642 NASLHFEPSKINIFHRDCK-----RSGRDATCLAAFLCTPIFLAPHFQITTV--GIRY 693  
 Db 619 VTLMSFSPAEPVHEVYECYSTSNKMKGVNII-----ICFQIKLYPQFQGLVANLYY 673  
 Qy 694 NATWDERRYTPRAHLDEGDRFTNRAVLLSSGGELCERINFHV--LDTADYVPVTSVY 752  
 Db 674 TLQLDGRH--TRRGLFPGRGHELRERNIAVTTMS--CTDFSFPFVCVQDLISPINVLSNF 731  
 Qy 753 SLEDPDHGP-----MLDDGWPTLTVS-----VPFWNGCNEDEHCVPLDLVLDARSDLPT 801  
 Db 732 SLWEESTPDQRAQKQDIPILRPSLHSTWHPFKXGCGDKKCEANL-----781  
 Qy 802 AMBYCORVLKPKQDCSAYTSLSDTTVFIIESTRQVAVATELNGENAYSTVLNISOS 861  
 Db 782 -----RVSPSPARSALRUTAPAS-----LSVELSLNLEEDDAYVYQLDLHFP 824  
 Qy 862 ANLQFASL--IQKEDSDGSIBC--VNEERLQKOV--CNVSYPPFRKAKAVAPRL-----910  
 Db 825 PGLSFRKVMELKPHSQIPVSCBELPEESRLLSALSCNVSSPIFKAGHVALQMENTLV 884  
 Qy 911 DSBFSKSIPLHLEIEILAAGSDNERDSTKEDNVAFLRFLHYEADV-----958  
 Db 885 NSSWGSDVELH-----ANVTENNEDSLLEDNATITPILYINILIQDEDSLYYS 938  
 Qy 959 FTRSSLSH-----YBVKLNS-----LERYDGIQPPFCIFRIONGLGPIPIKGM 1005  
 Db 939 FTPKGPKHQVXKHYVRIOPSIDHNIPTLEAVGVGPQ-----PSEG-----982  
 Qy 1006 KITIPATRSNRLKLRDLTDEANTSCNINWGNSTERYPTVE--BDLRRAPOLNHSND 1064  
 Db 983 -----PITHQ-----WSVQME--PPVPCYEDLERLPDAABPCLP 1015  
 Qy 1065 VVSINCMIRLVPOBINFHLLGNLW-----BSLKALKYKMKIMWNAALOROFSPFI 1118  
 Db 1016 GAFRCFV--VFRQELVQVIGTLELVGEIEASSMFSL--CSSLSISFNSS--KHFF--L 1067  
 Qy 1119 FREEDPSRQIEFEISQEDWQVPIWITVGSTLGLLLALLVLAALKLGF--RSARRRE 1177  
 Db 1068 YGSNASLAQVMKVVDVYVEKQ--LYLVLSIGIGLLLLLIIFVLVYGVFFKXNLKEME 1126  
 Qy 1178 PG 1179  
 Db 1127 AG 1128  
 RESULT 7  
 RWHLIC  
 cell surface glycoprotein CD11c precursor - human  
 N;Alternate names: leukocyte adhesion receptor p150,95 alpha chain  
 C;Species: Homo sapiens (man)  
 C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text\_change 22-Jun-1999  
 C;Accession: A36584; A36584; S00864  
 R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.  
 J. Biol. Chem. 265, 12750-12751, 1990  
 A;Reference number: A36584

A;Contents: erratum  
 A;Accession: A36584  
 A;Molecule type: DNA  
 A;Residues: 1-1163 <COR>  
 A;Note: this revision to the sequence from reference A35543 includes the carboxyl end  
 R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.  
 J. Biol. Chem. 265, 2782-2789, 1990  
 A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecu  
 A;Reference number: A35543; MUID:90153906; PMID:2303426  
 A;Accession: A35543  
 A;Molecule type: DNA  
 A;Residues: 1-834 <CO2>  
 A;Note: this sequence has been revised in reference A36584  
 R;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.  
 EMBO J. 6, 4023-4028, 1987  
 A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukoc  
 A;Reference number: S00864; MUID:88166645; PMID:3327687  
 A;Accession: S00864  
 A;Molecule type: mRNA  
 A;Residues: 1-755,'L',757-1163 <CO3>  
 A;Cross-references: GB:M81695; EMBL:Y00093; NID:9487829; PIDN:AAA59180.1; PID:9487830  
 A;Note: part of this sequence was confirmed by protein sequencing  
 C;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on  
 C;Genetics:  
 A;Gene: GDB:ITGAX; CD11C  
 A;Cross-references: GDB:119758; OMIM:151510  
 A;Map position: 16p11.2-16p11.2  
 C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat h  
 C;Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repe  
 F;1-13/Domain: signal sequence #status predicted <SIG>  
 F;20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>  
 F;120-1107/Domain: extracellular #status predicted <EXT>  
 F;149-319/Domain: von Willebrand factor type A repeat homology <VWA4>  
 F;1108-1133/Domain: transmembrane #status predicted <TM>  
 F;1134-1163/Domain: intracellular #status predicted <INT>  
 F;61.89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status

Query Match 18.2%; Score 1135.5; DB 1; Length 1163;  
 Best Local Similarity 28.4%; Pred. No. 2.4e-67;  
 Matches 354; Conservative 219; Mismatches 471; Indels 201; Gaps 48;

Qy 7 LVVAVALSIPGFTTFNMTRKPRVIPSRTAF-----FGYTVQOHDISGNKLVVGA 60  
 Db 10 LFTALATSL-----GFNLDTEE-----LTAFRVDSAGFGDSVQY---ANSVVVGA 53  
 Qy 61 PLETVGYOKTGVYKCPVIHGNCTKLNLRVTLNVSVSRKNNRMLGLSLATNPKDNSFLA 120  
 Db 54 POKITPAANTQGLYQCYSTGACBPGL-----QVPPEAVNMSLGLSLATSPSOLLA 107  
 Qy 121 CSPLWSHSCGSYYTTGMSRVNSNFRSKTVAPALQRC-QTYMDIVIVLDGNSIYP-- 177  
 Db 108 CGPTVHHCGRNWLTGLCLFLGPT-QLTQRLPVSRCQPCREQDIVLIDGSGISSEN 166  
 Qy 178 WVEVQHFLINILKKFYIGPGQIQGVGVQYGEDVVHFEHLNDYRSVKDWEAASHIEQ-RG 236  
 Db 167 FATMNFVRVAVISQFQ--RPSTQFELMQFSNKFQTHFTFEFRFRTSNPLSLIASVHQLQG 224  
 Qy 237 CTETETATGIEFARSEAFQKGRGAKKVMIVITDGESH--DSPDLKVIQOSERDNVTRY 295  
 Db 225 FTYTATATQNVVHRLFHASYGARRDATKILIVITDGKEGSLDKYKIVPMADAGIIRY 284  
 Qy 296 AVAV--LGYYNRGINPETFLNEIKYIASDPDPKRFVNTDEAALKDIDVALGDRIFSLEG 354  
 Db 285 AIGVGLAFQNRNS-----WKELNDIASKPSQEHFIKVEDFDALKDIQNLKEKIFAIEG 338  
 Qy 355 T-NKNETSFGLEMSQSGSSHVVEDGVLLGAVGAYDNGAVLKEYSACKVIP--LRESVL 411  
 Db 339 TETTSSSSFELEMAQEGFSVFTPDGVPVIGAVGVSFTWSGGAP-----LYPPNMSPTFI 391  
 Qy 412 KEFPEELKNHGAIFYTIVTSVVSSRQGRVYVAGAPRNFHTGKVLFTMNRSLTIHQAM 471  
 Db 392 NWSQENVDMDRDSYLGYS--TELALMKGVQSLVIGAPRYQHTGKAVIFT--QVSRQWMAEIV 449



QY 472 RQOIGSYGSSITSDIDGCVTDVLLVGVAPMYFNEGRCKVYVYELRQ--NRPVYNG 529  
 Db 450 TGTQGSYFGAGSLCSVDVDTGSDTLVLGAGHYVEQTR--GGQVSCVPLFRGWRWCD 508  
 QY 530 TLKDSHSYQARFGSSIASVRDLNDSYNDVVGAPLEDNHAGAIYIFHGFRG--SILKTP 588  
 Db 509 VLYGEGHGWGRGAALTVLGVDVNGDKLTDVVGAPGEEBNGAVYLFHVLGSPSH 568  
 QY 589 KQRTITASELATGLGYFGCSHGQDLNEDGLDLAVGALGNVILWSRPVQINASLHFE 648  
 Db 569 SORIASGQLSSRLQYFGQALSGQDLTDGLVDLAVGARGQVLLRLTRPVLVWGVSMQFI 628  
 QY 649 PKKINIFHRDCKRS--GRDATCLAAFLCF-----TFIAPLAFHQTITVIGIRYNAMDRR 701  
 Db 629 PAEIPRSARECEQVQVSEQLTQSNICLVYDKRSKNLLGSRDLQSS--VTLDLALDGR 685  
 QY 702 YTPRAHLDEGDRFTNRAVLLSSGQBELCERINFHLV---DTADYKVPFTSVEYSELEDP- 757  
 Db 686 LSPRAFQTKRSLSRVRLGL--KAHCE--NFNLLPSCVDSVTPTLRLNFTLVGPK 742  
 QY 758 ----DHGMDDGWPTTLRVSPVFWNGCNEDEHCVPLDVLVARDLPTAMEYCORVLRK 812  
 Db 743 LLAFLNRLPMLAADAQRYTASLPFEKNCADHIC----- 777  
 QY 813 PAQDSAYTSLPDTTFTTFTSTQRVAVATLENGENAYSTVLNISANLQFASLI-- 870  
 Db 778 --QDNLGISFSPGLKSLVGNLNAEVMVWMDGSDSYGTTITFSPAGLSYVAEG 835  
 QY 871 QKEDSDGSFECVNEERLQKQ-----VCNVSPFPRAKAKAERLDSBFS--KSIPLHLE 924  
 Db 836 QKQGLRSLHLCDSAPVGSQSTWSTCRINHLIFRGQAIFLATFVSPKAVLGDRL 895  
 QY 925 IBLAAGSDNERDSTKEDNAPLRLPHLYKAYEADVLTRSSLSHYEVLNNS----- 975  
 Db 896 LTAIVSSNNTPRTSK--TTFQLELFPVKY---AVTVWSSHEQFTKYNFSEBSEKSHV 950  
 QY 976 -LERYDIGIPFSCIFRIONGL--PPIHGMWKTITPATSNGNLLKLRLDPLDEANT 1032  
 Db 951 AMHY-----OVNLLGORDLV---SINFWPVVE-----LNQEA-- 981  
 QY 1033 SCNIWGN-----STEYRTPVEE---DLRAPQLNHSNDVVSINCLRLVP 1076  
 Db 982 ---VMDVEVSHQPNPNSLRCSEKIAAPPASDFLAHQKPNVLDCSIAGCLRPDC--VP 1035  
 QY 1077 N---QENFHLGN---WLRSKLKALKYSK---KIMVNAALQOFHSPTFRSDPSR 1126  
 Db 1036 SFSVQBELFTLKGNSLFGWVRQILQKQSVSVVAEITPDTSVYSQLPQGEAPRAQT 1095  
 QY 1127 QIR-FEISQEDQVPIWITVIGSLGGLLLALLVLAELKLGFFR 1170  
 Db 1096 VLEKXVHN-----PTPLIVGSSIGGLLLALLVLAELKLGFFR 1134

RESULT 8  
 S00551  
 leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse  
 N;Alternate names: complement-3 receptor alpha chain  
 C;Species: Mus musculus (house mouse)  
 C;Date: 30-Sep-1989 #sequence\_revision 30-Sep-1991 #text\_change 22-Oct-1999  
 C;Accession: S00551; 159078  
 R;Pytella, R.  
 EMBL J. 7, 1371-1378, 1988  
 A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the f  
 A;Reference number: S00551; MUID:88312584; PMID:3044779  
 A;Accession: S00551  
 A;Molecule type: DNA  
 A;Residues: 1-1153 <P>  
 A;Cross-references: EMBL:X07640; NID:g52982; PIDN:CAA30479.1; PID:g52983  
 A;Note: the authors translated the codon CAC for residue 569 as Glu  
 R;Sastre, L.; Roman, J.M.; Teplow, D.B.; Dreyer, W.J.; Gee, C.B.; Larson, R.S.; Roberts,  
 Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986  
 A;Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep  
 A;Reference number: 159078; MUID:86287312; PMID:2942940

A;Accession: 159078  
 A;Status: preliminary; translated from GB/EMBL/DBD  
 A;Molecule type: DNA  
 A;Residues: 11-44 <RES>  
 A;Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193  
 C;Genetics:  
 C;Gene: Mac-1  
 C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat h  
 C;Keywords: cell adhesion; glycoprotein; transmembrane protein  
 F;1-16/Domain: signal sequence #status predicted <SIG>  
 F;17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimen  
 F;148-318/Domain: von Willebrand factor type A repeat homology <WAA2>  
 F;1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 17.7%; Score 1104; DB 2; Length 1153;  
 Best Local Similarity 28.8%; Pred. No. 3e-65;  
 Matches 356; Conservative 202; Mismatches 506; Indels 170; Gaps 47;

QY 5 RGLVAVALSAMPQFTTFTMDTKPRVPGSRFAFFGTVQQSDHDSGNKMLVVCAPLET 64  
 Db 4 KALLIVT-ALALCHG-----FNLDTSEHPTFQENAKG-FQONVYQ---LGGTSVVAAPQEA 54  
 QY 65 NGYKTDGVKCPVHGNCTKLNLRVTLNSVSRKNNMRLGLSLATPKDNSPLACSP 124  
 Db 55 KAVNQTGALYQCDYSTSRCHPIPL-----QVPPEAVNMSLGLSLAVSTVQQLLACGPT 108  
 QY 125 WSHCGSYTYTGMCSRYSN--FRPSKTVAPALQRC-QTYMDIVIVLDGNSI--YPWVE 180  
 Db 109 VHONCKENTYVNGLCYLPGLSGNLLRPQPPALRECPQESDIVFLIDSGSINNIDPQK 168  
 QY 181 VQFLINILKKFYIGPGQIQGVYQYGEDVHVEHFLNDYRSVKDVEAAHSEORGGTET 240  
 Db 169 MKEFVSTVMEQF--KSKTFLPMLQYSDFFRIHFTFNDPKRNPSPRSVHSIKQLNG-RT 225  
 QY 241 RTAFGIEPARSEAFQK--GGRKGAKVMIVITDGSBDSHSP--DLEKVIQSSRDNVTRYAV 297  
 Db 226 KTASGIRKVVRELHFKTNGARENAKILVITDGEKPGDPLDYKQVPEADRAGVIRVI 285  
 QY 298 AVLGYNRRGINPTFLNELIKYIASDDDDKXFFVNTDEALKDIDVADGDFISLGGTNK 357  
 Db 286 GVGNAFNK---POS--REELDTIASPAGEHVPQDNPEALNTIQNQLEKIFAJEGTQ 340  
 QY 358 NET--SPGLEMSQTFSGSHVEDVGLLGAAGVAYMNGAVLKETSAGKVIPLRESYLKBPPE 416  
 Db 341 GSTSFSFHEMSQEFASITNSGFLASVSGSPDAGGAFLYTSKDKV-----TPINTRV 395  
 QY 417 ELKNHAYLGVTVTVSVSSRQGRVYVAGAPRNFHTGKVLITMNNRLSLTHQMRGQOI 476  
 Db 396 DSDMDAYLGY--ASAVILNRVQSLVGLGAPRYQHIGLVWMP--RENFGTWEPHTSIKSGOI 453  
 QY 477 GSYFSEITSDIDGCVTDVLLVGVAPMYFNEGRCKVYVYELRQ--RFPVYNGTLKDS 534  
 Db 454 GSYFASLCSVDMDADGNTNLLIGAPHYIEKTR--GGQVSCVPLFRGWRWCD 512  
 QY 535 HSYQVAFSGSSIASVRDLNDSYNDVVGAPLEDNHAGAIYIFHGFR--GSILKTPKQRT 593  
 Db 513 QGHFWGRFGAALTVLGVDVNGDKLTDVAIGAPGEQENQAVYIFYGASIASLSASHSHRI 572  
 QY 594 ASELATGLGYFGCSHGQDLNEDGLDLAVGALGNVILWSRPVQINASLHFEPSKI- 652  
 Db 573 GAHFSPLGQYFGQSLSGGKDLTMDGLMDLAVGAQHLLLRQAQPVIRUEATWESPKVA 632  
 QY 653 -NIFHRDC-----KRSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMBERRYP 704  
 Db 633 RSVF--ACQEQVLNKKDAGEVRVCLVRKNTKDLREGDIQST---VTYDLALDPRGRI 687  
 QY 705 RAHLDEGGDRFTNRAVLLSSGQBELCERINFVLDTA--DYKVPFTSVEYSL--EOP----- 757  
 Db 688 RAPFDETKNN--TRRTQVFGLNQKCEKTLKGLTLPQVDSDVSPILRLNLTIVLGSPLRFG 746  
 QY 758 DHGPMDDGWPTTLRVSPVFWNGCNEDEHCVPLDVLVARDLPTAMEYCORVLRKPAQDC 817  
 Db 747 NLRPVLANDQAQRYTAMFPFPEKNCNDSICQDDLSI----- 782

A;Cross-references: GB:M18044  
A;Hickstein,D.D.; Hickey,M.J.; Ozols,J.; Baker,D.M.; Back,A.L.; Roth,G.J.  
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989  
A;Title: cDNA sequence for the alpha subunit of the human neutrophil adherence recep  
A;Reference number: A32218; MUID:89098893; PMID:2563162  
A;Accession: A32218  
A;Molecule type: mRNA  
A;Residues: 9-1153 <HC>  
A;Cross-references: GB:J04145; NID:g189068; PIDN:AAA59903.1; PID:g386975  
A;Note: part of this sequence was confirmed by protein sequencing  
R;Fleming,J.C.; Pahl,H.L.; Gonzalez,D.A.; Smith,T.F.; Tenen,D.G.  
J. Immunol. 150, 480-490, 1993  
A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha  
n during evolution.  
A;Reference number: A46526; MUID:93123748; PMID:8419480  
A;Accession: A46526  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-499,501-1153 <FILE>  
A;Cross-references: GB:S52227; NID:g263047; PIDN:AAB24821.1; PID:g263049  
A;Note: the last three bases of intron I3, CAG, are included in some but not all manu  
A;Note: sequence extracted from NCBI backbone (NCBIP:121963)  
R;Pierce,M.W.; Remold-O'Donnell,E.; Todd III,R.F.; Arnaout,M.A.  
Biochim. Biophys. Acta 874, 368-371, 1986  
A;Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across  
A;Reference number: A90664; MUID:87076671; PMID:3539202  
A;Accession: A26091  
A;Molecule type: protein  
A;Residues: 17-31 <PE>  
A;Experimental source: granulocytes  
R;Pahl,H.L.; Rosmarin,A.G.; Tenen,D.G.  
Blood 79, 865-870, 1992  
A;Title: Characterization of the myeloid-specific CD11b promoter.  
A;Reference number: I52567; MUID:92144986; PMID:1346576  
A;Accession: I52567  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-9 <RES>  
A;Cross-references: GB:M84477; NID:g180184; PIDN:AAA51960.1; PID:g553219  
C;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac  
C;Genetics:  
A;Gene: GDB:ITGAM: CR3A  
A;Cross-references: GDB:120599; OMIM:120980  
A;Map position: l6pl1.2-l6pll1.2  
A;Note: promotor contains a GATA motif and two Spl consensus binding sites  
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat h  
C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer;  
F;1-16/Domain: signal sequence #status predicted <SIG>  
F;17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>  
F;17-1108/Domain: extracellular #status predicted <EXT>  
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>  
F;465-473/Region: calcium/magnesium binding #status predicted  
F;530-538/Region: calcium/magnesium binding #status predicted  
F;593-601/Region: calcium/magnesium binding #status predicted  
F;1109-1134/Domain: transmembrane #status predicted <TM>  
F;1135-1153/Domain: intracellular #status predicted <INT>  
F;86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Bind

|    |                       |   |
|----|-----------------------|---|
|    | Query Match           | 17.6%; Score 1094; DB 1; Length 1153;                             |
|    | Best Local Similarity | 28.6%; Pred.No. 1.4e-64;  |
|    | Matches               | Conservative 208; Mismatches 517; Indels 158; Gaps 46;            |
| Qy | 12                    | ALSLWPGFTDTFMNDTRKPRVIPGSRTAFGYTVQQHDISGNKLWVGAPLENTNGYQKTG 71    |
| Dd | 10                    | AULTCHG----FMLDTEANMTQEENARG-FQGSSVQ--LQGS-RVVVGAPQEIVAANQRG 61   |
| Qy | 72                    | DVTYKCPVIHGCTKNLGMGVTTLSNVSKDKNNMLGLSLATNPKDNSFLACSPWLWSHECGS 131 |
| Dd | 62                    | SLYQCXYSTGCSEPIRL-----QPVEAVNMASLGSLAAATSPPOLLAGCGPTVHQTCSE 115   |
| Qy | 132                   | SYTWTGCGSRVNSENFSXTVP-ALQRC-QTYMDIVILVDGNSNIYP--WWBVQHFLIN 187    |
| Dd | 116                   | NITYVKGLCFUFGNLRIQQCKPFELARGCPQEDSDIADILDGSGSIHPDFRMKEPVST 175    |





QY 686 -TTVTGIRYNATNDRRYTPRAHLDGEGDRFTNRAVLSSGOELCERINPHV-LDPADV 743  
Db 663 GRLLAMSLYTLQDGHMRSGFLFPDGSHELSCNTSI--TPDKSLDFFHFPICQDLI 720  
QY 744 KPVTSVEYSLEDDHGMDDG-----WPT--TLRVSPFMNCGNDEHCVPLDLD 794  
Db 721 SPINVLNLSLEEGTPTDQKRAMQPIRPSIHVTKEIPEKXCGEDKCEBANLTL 780  
QY 795 --ARSOLPTAMEYQORVLRPAQDCSAYTSLPFTTFFIIESTQRVAVATLENGENAY 852  
Db 781 SPARSG-----PLRLMSSASL-----AVBWTLSNGEDAY 810  
QY 853 STVLNISQSANTLOF--ASLIQ-----KEDSDGSIECVNEERLOQOV-CNVSYPEF 900  
Db 811 WVRLODFPRGLSFRKVEMLQPHSRMPVSCBELTEGS-----SLTTLKCNVSSPIF 863  
QY 901 RAKAKVAFR-----LDSEFSKSIPLH--HLBIELAAGSDSNEDSTKEDNVAPLPHL 951  
Db 864 KAGQSVLQWENTLNSWEDFVNLGTVHC-----NENSSLQEDNSAATHIPV 914  
QY 952 KYEADVL-----PT-----RSSLSH-YEVKLNS-----LRYDGIQPPF 986  
Db 915 LYPVNLTKQENSTLYISPTPKPKTQQQVHYQRIOPSAYDHNMPTLEALVGVPRPH 974  
QY 987 SCIFRIQNLGLPPIHGMMKTIPIATRSNRLKLRLDFTDEANTSCNMGNSTEYRPT 1046  
Db 975 S-----EDLIT-----YTWVSQTDPLVT 992  
QY 1047 PVTEDLRAPQLNHSNDVVSINCLRLVNPQINPHLQNLWL-RSLKALKYSM--KI 1103  
Db 993 CHSEDLKR-PSSBARQCLPGVQVCPVFRWEILQVGTVELSKEIKASSTLSLCSL 1051  
QY 1104 MVNAALQROPHSPPIFREEDPSQIIEFISKQEDQWQVPIWIVGSLGGLLILALVLAL 1163  
Db 1052 SVSFSKSKPH--LVGSKASEQVLVKVDLIHEKEM-LHVYVLSGIGGLVLLFLFLAL 1107  
QY 1164 RXLGFPP-RSARRRR 1177  
Db 1108 YKVGFFKRLKRW 1122

RESULT 11  
A53213  
Integrin alpha-B chain - human  
C;Species: Homo sapiens (man)  
C;Date: 19-Oct-1995 #sequence\_revision 31-Mar-2001 #text\_change 31-Mar-2001  
C;Accession: A53213  
J;Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.  
J;Bio. Chem. 269, 6016-6025, 1994  
A;Title: Molecular cloning of the human mucosal lymphocyte integrin alpha (B) subunit. Un  
A;Reference number: A53213; MUID:94164962; PMID:8119947  
A;Accession: A53213  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1179 <SHA>  
A;Cross-references: GB:L25851; NID:9457244; PID:9457245  
C;Genetics:  
A;Gene: ITGAB  
A;Cross-references: GDB:330801  
A;Map position: 17p13  
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
F;193-371/domain: von Willebrand factor type A repeat homology <VMA3>

Query Match 14.98; Score 924.5; DB 2; Length 1179;  
Best Local Similarity 26.3%; Pred. No. 3.2e-53;  
Matches 338; Conservative 206; Mismatches 482; Indels 259; Gaps 52;

QY 23 FNMDTKPRVIPSRTAFFGVYQOQHDISGNK-WLVVGAPLETNGYQKTGDVYKCPVTEG 81  
Db 19 FNVDAVAPLTPKGAFFVLSLLHQDPSTNQWLLVTSF--RTKRTGRLRCSLQVD 75  
QY 82 NCTKLNLGRVTLNVSERKQNRGLGLSLATNPKNSEFLACPLW---SHECGSYTTGM 138

Db 76 BILCHPVEHVPIPKRRHR-----GVTVVRG--HHGVLCIQVLVRRPHSLSSE--LTGT 125  
QY 139 CS-----RVNSNFRFSK-----TVAPALQRC----- 159  
Db 126 CSLLPDLRPOAQANFPLENLDPDARVDGDCYSNKEGGEDDVTARQRRALKEBE 185  
QY 160 -----QYMDIVILDCSNSTYP--WVVOHFLNLLKKFYIGQIQGVGVQ 205  
Db 186 EKEBEEDBEERAGTBIAIILDGSGSIDPDPFORAKDFISMMNRNFYEKFECPNFALVQ 245  
QY 206 YGEDVVFHFLNDYRSVKDVVEAASHIEQRTGTETRTAFGIEFARSEAF--QKGRKGAK 263  
Db 246 YGGVQTQFOLDSDQVWASLARVQNTQVGSV-TKTASAMQHVLDSIFTSHSGSRKAS 304  
QY 264 KMIVITDGSHEP-DLEKVIQOSERNVTRVAVAVLGYNNRGINPPT-----LNEI 317  
Db 305 KMVVLTDGGIFEDPLNLTITVINGPOMQGVREPAIGV-----GEEPKSARTAREL 354  
QY 318 KYIASDPDDKHFNFVNTDEAALKDIVDALGDRIFSLGNTNKNQETSGLMSGTGFSHVVE 377  
Db 355 NLIASDPDETHAFKVTNTMALDGLLSKLRNLIISMEGTVGDLHY--QLAQIGPSAQILD 412  
QY 378 D-GVLLGAVGAYDWN-GAVLKETSAGKVIPLRESYLKEPPELKNHGAYLGYTTSVSS 435  
Db 413 ERQVLLGAVGAPDMGGGALLYDTRRRGRFLNQTAADAAADAAQY-SYLGAVV-AVLHK 470  
QY 436 RQGRVYVAGAPRFNHTGKVIILFTMHNRSLTIHQAMRQOQIGSYFGSEITSVDIDGDGVT 495  
Db 471 TCLSLYVAGAPQYKHEGAVPEL-QKEGREAFLPVLEGEOMGSIYFGSLCPVDIDMDGST 529  
QY 496 DVLLGAPMYFNEGRERKQVYVEI--RQNPVYNGTUKDSHYONAPFGSIAVRDLN 553  
Db 530 DFLVAAAPFYVHG--REGVTVYVRLSEQDGSFSLARILSGHPGFTNAPFGFMAAAGDLS 588  
QY 554 QDSYNDVVVGAPLE---DNHA--GAIYIFHGRGSIILTKPKORITASELALGCVPGCS 607  
Db 589 QDKLTDVAIGAPLEGAGDAGSFGSVIYNGHWDGLSASQSRTRATVAPGLQYFGNS 648  
QY 608 IHGQDLNEDGLIDLAVGALGNVILMSRPVQVNASLHFEPSKINIHRDCKRSGRDT 667  
Db 649 MAGGFDISGDGLADITVGTLGQAVVFRSRPVRLKVSMAFTPSALPIGNG----- 699  
QY 668 CLAAELCTPIPLAHPFQTTVGIR---YNATMDERRYTPRAHLD-----E 710  
Db 700 VNVNLCPE---ISSVTASGSLREALNFTLDVDVGKRRRLQCSVRSCIGCLRWS 756  
QY 711 QGDRFTNRAVLSSQOELCERINPHVLTADYVKPVTFSVEYSLEDP----DH-GPMLDD 765  
Db 757 SCSQLCEDLLMPTGELCEB-----DCFSNASVKSVYQLQTPESQTDHPQPIldr 807  
QY 766 GHPFTLRVSPWNGCNDEHCVPLVDARSDLPTAMEYCORVLRKPAQDCSAYTSLSPD 825  
Db 808 YTEPAIFQLPYEKACKNKLFCVABLQ-----ATTVSQO 842  
QY 826 TTVFIIESTRQVAVEATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNBE 885  
Db 843 E---LVVGLTKELTLNINLNSGDSYNTSMALNYPRLQLKEM-QKPPSP-NIQCDDPQ 897  
QY 886 RRLQKQV--CNVSYPPFR---AKAVAPRLDSEFSKSIPLHLELELAAGSDSNEDSTK 940  
Db 898 PVASVLNMCRIHGVLPKXSSAHVSVVWQLE-----ENAPPNRTADITVTVINSNRRSLA 953  
QY 941 EDNVAPLRFHLKVEA-----DVLFTRS--SSLSHYEVKLNSLSELYDGIQPPFCIFRION 994  
Db 954 NE-THTLQPRHGFVAVLSKPSIMYNTQGLSHKE----- 988  
QY 995 LGLFPIHG-----MMKTIPIATRSNRLKLRLDFTDEANTSCNMG--NSTEYRP 1045  
Db 989 -FLFVHGENLRGABYQICVPTKLR-GLQVAAVKLRTQASTVCT--MSOERACAYSS 1045  
QY 1046 TPVEEDLRAPQLNHSNDVVSINCLRLVNPQEI-----NPHLLGNLWLSRLKALKYKS 1100  
Db 1046 VORVBEHWSVCVIAADKENVTVAAEISNDHSEBELKDVTELQILQELISFN-----KS 1098

|    |      |                   |               |             |                |                          |
|----|------|-------------------|---------------|-------------|----------------|--------------------------|
| Qy | 1101 | KKIMVNAALQRPQSPFI | FRBEPDPSQIBEI | SKQEDWQVPTI | WNGSTGGHLLALLV | 1160                     |
|    |      |                   |               |             |                |                          |
| Db | 1099 | LVEGLNA           | -----         | ENHRKTI     | VTVFKDEKIH     | SLPIIINGSVGGLLVIVIL 1143 |
|    |      |                   |               |             |                |                          |
| Qy | 1161 | LALRKLGFPRARRR    | REPCGLDTPK    | 1185        |                |                          |
|    |      |                   |               |             |                |                          |
| Db | 1144 | VILFKCGFFK        | -RKYQLALE     | SIRK        | 1166           |                          |
|    |      |                   |               |             |                |                          |

## RESULT 12

I58409  
 integrin alpha-9 chain precursor - human  
 CSpecies: Homo sapiens (man)  
 CDate: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 31-Dec-2000  
 CAccession: I58409; A49459  
 R:Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takahashi, K.  
 Oncogene 9, 611-619, 1994  
 A>Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in small cell lung carcinoma  
 A.Reference number: I58409; MUID:94119603; PMID:8290272  
 A.Accession: I58409  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A.Molecule type: mRNA  
 A.Residues: 1-1035 <RES>  
 A.Cross-references: GB:D25303; NID:5464180; PIDN:BAA04984.1; PID:G533327  
 R:Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytel, R.; Sheppard, D.  
 J. Cell Biol. 123, 1289-1297, 1993  
 A>Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel partner of the alpha5 beta1 integrin  
 A.Reference number: A49459; MUID:94064789; PMID:8245132  
 A.Accession: A49459  
 A>Status: preliminary  
 A.Molecule type: mRNA  
 A.Residues: 30-1035 <PAL>  
 A.Cross-references: GB:I24158  
 C:Superfamily: integrin alpha-4 chain  
 CKeywords: glycoprotein; metal binding; transmembrane protein  
 F1:1-27/Domain: signal sequence predicted <SIG>

|                       |              |                  |                 |              |
|-----------------------|--------------|------------------|-----------------|--------------|
| Query Match           | 12.8%        | Score 799.5;     | DB '2;          | Length 1035; |
| Best Local Similarity | 24.2%;       | Pred. No. 6e-45; |                 |              |
| Matches 307;          | Conservative | 177;             | Mismatches 418; | Indels 369;  |
|                       |              |                  |                 | Gaps 49;     |

|     |    |     |   |     |
|-----|----|-----|---|-----|
| 5   | QY | 5   | RLVWAWALSLWPGFTDTFNKOTRKRPRVTPGSRATAFGYTVQOQHDISGNKWLVLVGAPLET  | 64  |
| 14  | Db | 14  | RALLLALVVGATIP--AGAYNLAPQRPVHFQGPADSPFGYAVLEHFFHFDNTRWLVVGAPKAD | 71  |
| 65  | QY | 65  | NGV-----QKTGVYKCPVITHGN---CTXLNLGRVTLSNVS-----ERKDNMRIGLSLA     | 110 |
| 72  | Db | 72  | SKTSPSVKSPGAVFKCRV-HTVNPDRCTELDMARGKNRGTSCKTCREDRDDSEWGSLSLA    | 130 |
| 111 | QY | 111 | TNPK-DNSFLACPLMSHECGSSYYTT-----GMC SRVNSNFRP-SKTVPAPALORCQTY    | 162 |
| 131 | Db | 131 | RQPKADGRVLACAHRWK-----NIYYEADHILPHGFVCYIIPSNLQAKGRTLIPCYEB----  | 182 |
| 163 | QY | 163 | NDIVIVLDCSNSIYPWVEVQHFNLINTLKFEYIGPGQIQGVVQYGEDVHVEPHLNDYRSV    | 222 |
| 183 | Db | 183 | -----YKKKY-----   | 187 |
| 223 | QY | 223 | KDWEAASHIEQRGGTETRTAFGIEFARSEAFQKGRKGAKKVMIVITDGEHSDSPLEK       | 282 |
| 188 | Db | 188 | -----GEEHGS-----  | 193 |
| 283 | QY | 283 | VIQOSERDNTRYAVAVILGYNNRGINPETFLEINIKYIASDDPKHFFWVTDEAALKDVI     | 342 |
| 194 | Db | 194 | -----   | 193 |
| 343 | QY | 343 | DALGDRIFSLEGTNKNETSFGLEMSOTGFSSSHVVVZDGVLLGAVGAYDMNGAVLKETSAKG  | 402 |
| 194 | Db | 194 | -----CQAGIAGFPTEELVWNGAPSGSYWAGTI-----K                         | 222 |
| 403 | QY | 403 | VIPLRE-SYLKEFPPEELKNHG-AYLGYTVTSVVSRSQGRV-YVAGAPRNRHGTGVILFTM   | 459 |
| 223 | Db | 223 | VLMLETDNTYKLNDEYIMNRRVTYLGYAVTAGHFHSHPSTIDVVWGGAPQDKIGIKVYIFRA  | 282 |

|      |    |   |
|------|----|---|
| 460  | QY | HNRS---LTHQAMRGQOIGSYFOSSITSVDIDGCVYTDVLLVGAHPYFNEGRGKVY 516      |
| 283  | Db | DRRSOTLIIKIFQA-SGKMGSGYFOSSLCVADNLNGLSD-LLVGAHP-FSEIRDSQVT 338    |
| 517  | QY | VYZELONRFVYNGTLKDSSHSONASFGSSIASVRDLNODSYNDVVVGAPLEDNHGAIAIY 576  |
| 339  | Db | VYINRGNGALEBQLALTGDGAYNAHFSGSIASDLDLNDGFPDVAIGAPKEDDFAGNVI 398    |
| 577  | QY | FHGFSGSILKTPKQRIITASELATGUYFGCSITHGOLDNEDGLIDLAVGL--GNNAVILW 634  |
| 399  | Db | YHGDAGGIVPOYSMKLSGKINPILVRMFQSGISGGIDMDGNGYDPDVTVCFAFMSDSVLLR 458 |
| 635  | QY | SPPVQIYNASLHFPBSKINIHRDCKRGRDATTCLAAFLCPTPIFLAHPHFOITTVGIRN 694   |
| 459  | Db | ASPVITVDSI-FLFGSINITAPOCHDGOQPVNCLNVTTCFS--FHGRHV-PBEIGLNVY 514   |
| 695  | QY | ATMD---ERRYTPEAH---LDEGDRFTNRVALLSSGOELCERINPHV-LDTADYVKDV 746    |
| 515  | Db | LMADVAKEKGOMPRVYFVLGSTMQVTEK-LQTYMEETCRHYVAHVKKRVQDVISPI 573      |
| 747  | QY | TFSVYSYSLD-----PDHGPML--DGMPTTLRVSPVFNMGNCNDEHCVDPLVLD 794        |
| 574  | Db | VPEAAYSLSHVTEGEERELPPLTVLAWKKGOKLAQKNQTVFERNC-RSEDCAADLQLO 632    |
| 795  | QY | ARSDLPANMEYQORVLRFPAQDCSAYTILSPDTTVILIESTRQVAVEATLENRGENAYST 854  |
| 633  | Db | GKL-LLSMD-----EKTLYALGAVXNISLNTSISNLGDDAYDA 671                   |
| 855  | QY | VLNISQSANLOPASLIQKEDSGSIECVNEERRLOQVCNVSYPPFRKAKAVAFRLDSRF 914    |
| 672  | Db | NVSFNVSRELPIFNWQKEEM--GISCELLESDFLK--CSVGFPEMRSKY-----EF 721      |
| 915  | QY | SKSIFLHLEIY-----LAAGSDNBRDSTYKENVAPLRHLKYEADVLT--RSSSL 965        |
| 722  | Db | SVIFDTSLSGSEEVLSPIVTAQSNTERSESLHDNTILVMLPHEVDTSITGINSPTS 781      |
| 966  | QY | SHYEVKLASS---LERYDGIQPPSCIFRIQNLGLFPIHGMMKTIPIATRSQN-RLL 1020     |
| 782  | Db | FYGESVDAANPIQLDDLECHPOPINILOYNTGSPSTLPGSSVSGISFNRLSSGGAEMP 841    |
| 1021 | QY | KLRDLTDEANTSCNWNSTYEYRTPV-----EEDL-----RRAPOLNHSNS 1063           |
| 842  | Db | HVQEMVVGQEKNGCSFOKN-----PTPCIIPOEQENIFHTPIAFFTKSGRKVLCDCKPGI 896  |
| 1064 | QY | DVWSINCMIRLVPQOE---INFHLGNLWLSLKALKYKSMKIMYNAALOROFHSPPIFR 1120   |
| 897  | Db | SCLTAHCNFSALAKEBSRTIWIYMLLN-----TEILKXKSSSVI-----QFNGRKVKY 944    |
| 1121 | QY | REDPSRQIETFEITSKOEDMQVPI-----WIIVGSTIGGLLILALLVLALRKL 1166        |
| 945  | Db | VDPALRVVYIAHGNPEEYTVFEALHLEPRGYVVGWIIAISLIWGLIFLLAVILWQW 1004     |
| 1167 | QY | GFRSARRRRR 1177   |
| 1005 | Db | GFRF--RRYKE 1013  |

## RESULT 13

A55348  
 integrin alpha-1 - chicken (fragment)  
 C-Species: Gallus gallus (chicken)  
 C-Date: 10-Feb-1995 #sequence\_revision 10-Feb-1995 #text\_change 15-Sep-2003  
 C-Accession: A55348  
 R:Kern, A.; Briesewitz, R.; Bank, I.; Marcantonio, E.E.  
 J. Biol. Chem. 269, 22811-22816, 1994  
 A>Title: The role of the I domain in ligand binding of the human integrin alpha-1beta-  
 A-Reference number: A55348; MUID:94357930; PMID:7521332  
 A-Accession: A55348  
 A>Status: preliminary  
 A-Molecule type: mRNA  
 A-Residues: 1-272 <KER>  
 A-Cross-references: GB:U10114

F;55-230/Domain: von Willebrand factor type A repeat homology <WNA2>

Query Match 12.7%; Score 789; DB 2; Length 272;  
Best Local Similarity 55.9%; Pred. No. 3.2e-45;  
Matches 151; Conservative 53; Mismatches 62; Indels 4; Gaps 3;

Qy 108 SLATNPKDNLPLACSLPWSHECGSSVYTTGMCVRVNSNFRSTKTVAPALQRCOTYMDIVI 167  
Db 2 TLVTPK-GGLACGLPIAYKQGLHVTTCVNSVSTFTVAVAPSVQECQTQDVI 60

Qy 168 VLDGNSIYPWVQVHPLINILKKFYIGPQIQGVVQVGEDVVFHFLNDYRSVKDVE 227  
Db 61 VLDGNSIYPWESVTFALNLLNMDIGPQQTQGVIGVQGVVHFLNYSTTBEVMD 120

Qy 228 AASHIEQGTETRTAFGIFPASEAFQK--GEGKAKKVMIVITDGHSDSDIAKVIQ 285  
Db 121 ALRIRGGTQTALGIDTAREAPTEAHGARGVQKVMIVITDGHSDNTRLOEVID 180

Qy 286 QSRDNRVRYAVVLGYNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIDVAL 345  
Db 181 KCDENIFQAIALLGSYRGNLSTKFEVEIKSIASKPKTEKFFNVSDALAVTIVEAL 240

Qy 346 GDRIFSLLEG-TNKNSTSPGLEMSQTFSSH 374  
Db 241 GERIFALEATTQQAASFEMEMSOAGPSAH 270

RESULT 14  
T31437  
integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)  
C;Species: Lytechinus variegatus (variegated urchin)  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jan-2000  
C;Accession: T31437  
R;Hertzler, P.L.; McClay, D.R.  
submitted to the EMBL Data Library, May 1998  
A;Description: Alpha SU2, a sea urchin integrin which binds laminin.  
A;Reference number: 221035  
A;Accession: T31437  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1041 <HER>  
A;Cross-references: EMBL:AF067658; NID:G3220240; PID:G3220241; PIDN:AAC23572.1  
A;Experimental source: developmental stage embryo  
C;Function:  
A;Description: binds laminin  
C;Superfamily: integrin alpha-2b chain

Query Match 11.9%; Score 738; DB 2; Length 1041;  
Best Local Similarity 23.7%; Pred. No. 7.9e-41;  
Matches 308; Conservative 168; Mismatches 422; Indels 402; Gaps 51;

Qy 3 LPRGLVAVALLSLPFGTDTFNDMTKPRVPIGSRTPAFPGYTVQOHDISGNKWLVTGAP- 61  
Db 1 MARILLSPVAILLDSTAGFNLDLAPLKFDPGDSMGFSAQHRDGNWDNALVGAPE 60

Qy 62 LETN-GVQKTGDVYK-CPVTHGN--CTKLNLGRV-TLSNVSEKKNRMLGLSLATNPKD 115  
Db 61 ASTNQSGVYTKGGAIVYVCRPLSGPSVCEQIPFDKTNINNV--KSNQWFGATVASGGAN 118

Qy 116 NSFACSP--LWSHECGSSVYTTGMCVRVNSNFRSTKTVAPALQRCOTYMDIIVILGDSN 173  
Db 119 GQILACAPRLVW----- 130

Qy 174 SIYPWVEVQHPLINILKKFYIGPQIQGVVQVGEDVVFHFLNDYRSVKDVEAASHIE 233  
Db 131 -----VETASVTM 138

Qy 234 QRGGTETRTAFGIFPASEAFQKGRKGAKKVMIVITDGHSDSPDLEKVIQOOSERNVT 293  
Db 139 ROEREPTGTCT----- 149

Qy 294 RYAVAVLYGNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSL 353

RESULT 15

JC7294

alpha integrin - sea urchin (Strongylocentrotus purpuratus)

C;Species: Strongylocentrotus purpuratus (purple urchin)

C;Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 02-Sep-2000

C;Accession: JC7294

Db 150 -----VGDSDFTNFVNS-----PCQ 165

Qy 354 GTNKNETSFG-LEMSQTCFSSHVVED--GVLLGAVGAYDMGAVLKETSAGKVIPLRESY 410

Db 166 STNREDYDFDKITHCQAGFSGIIFSDNSALVWAPGSIYLOQGIYVQSL-----LNSV 219

Qy 411 LKEFPPEELKNHG-----AYLGTYTTSVVSSRQG-RVTVAGAPRPNH-TGKVLFTMHN 462

Db 220 VQATOB--SNTGYTSFDSYRGYSIALGDFNGDGVDVVGTPRAESLMGLVAIPDQNLN 277

Qy 463 RELTTHQAMRGQOIGSYFGSEITSVDIDGDGVDVLLVLCAPMYFNEGR-----EGKVVV 517

Db 278 Q---FNQVM-GTQIVAVFGYSVTVDINDTIVDD--LLVGAPMYMDGPAIQWSEAGAVV 332

Qy 518 YELR-----QNRFFVNGTLKDSHSYQNAFGSSIASVVDLNDQSDYNDVVVGAPLEDNH 570

Db 333 YLQNPDVPGASNRLLSSTLLIGQI--RSRKLASIASIGSNQDQDFNDVAIGAPEGDD 390

Qy 571 AGAIVYFHGRGSIILKTPKQRTASELA-TGLQYFGCSIHGQOLDLNEGLDLAVGA--L 627

Db 391 AGAIVYHGSANGLASTPAQVLTPSTGLHSGITTFGLQGGQDMKRYPOLLYGABSA 450

Qy 628 GNAVILMSRPVQVINASLHFEPSKINIFHRDCKRGRDATCLAAFLCTPIFLAHPFOT 687

Db 451 NTAVLIRTPVVSLLDALTNTBFIGINLENKIYELA--DGTWTSFIAMTCFTYTGNYLPD 508

Qy 688 TVGIRYNATMDERRYTPR--AHLDRGDRFTNRVLLSSGQELCERINPHVLOT-ADYVK 744

Db 509 HIDISVTVTDSGLIANRRAMEVDMDSEITKTRELAVSTQ-FCDFLRAVYVNGSIEDKLT 567

Qy 745 PVTFSVEVSELDPDHG-----PMLDDGWPITLAVSVFPWNGCNEDEHCVDPOLVDARS 797

Db 568 PIKVLQVLDLNDSESLRLOPHBILPIIDMATWSTQTKQVSIQNNC-VNNICIPOLDVTVTP 826

Qy 798 DLPTAMEYQCVLRKPAQDCSAYTSLFDTTVEIIBSTRQVAVATELENRGENAYSTVLN 857

Db 627 NLPN-----IVIGTQOELTLDVSLNRRGSDAFQSSLS 658

Qy 858 ISQSNLQFASLIQKEDSDGSEICVNEERLQKQCVNTSYPEFRKAKVAFRLD-SEPSK 916

Db 659 VTYPIGLQFVRLERKANMDFSVTC-SESDSLRIITCDTGNPMV-GRNILEFGLTJSTPQV 716

Qy 917 SIFLHLEITELAAAGSDSNRSDTKEDNVAPLAFHLKYEADVLTFTSSSSISHYEVKLNSSL 976

Db 717 SGDKDSIEFPYKAESENSSDPNTLENNELNMTVPVTVDCTL--KLLSASYEIVMYSTQ 773

Qy 977 ERYDGIQPPFSC-----IFRIQMLGLFPFHGMKMITIPIATSGNRLK 1022

Db 774 EDY--VVPFPKAXNASEADIGMVMHLYEVRNTGSSNAGEVSLNQWPKQNEDEYLFYL 831

Qy 1023 RDLTDEAMTSCNI-----WGNSTERYPTPVEEDLRRAPQLNH--- 1060

Db 832 LGIMTEGVT-COLYQKAPNPSGVKLEPSTKAKLSNT---TQVSGRKRREPEVAEALA 886

Qy 1061 -----SNSDVVINCINIRLVPMQETN-----PHLLGNLWLSLK----- 1094

Db 887 QTDNVITYCASDSCVLINCTI-----DEINASKKVVRILGRFWERTFOKAVSELTPVQQA 941

Qy 1095 -----ALKYKMKIMVNAALQRFH-----SPPIFEE--DPSRQTEPEISKQEDMQ 1139

Db 942 TIASSASAAVK---TIPYNILPRDPSDSTKASTLVTEBLVPPVTPPIA-----W- 988

Qy 1140 VPIWIVYTLGGLLLALLVLAALRKLGFPRARRRREP 1179

Db 989 ---WIIVSVLGGIILLIILGLWKCQFP-----BRKKPG 1021

R: Susan, J.M.; Just, M.L.; Lennarz, W.J.  
Biochem. Biophys. Res. Commun. 272, 929-935, 2000  
A: Title: Cloning and characterization of alphaP integrin in embryos of the sea urchin Sc  
A: Reference number: JC7294  
A: Contents: Embryo  
A: Accession: JC7294  
A: Molecule type: mRNA  
A: Residues: 1-1054 <SUS>  
A: Cross-references: GB: AAD55724  
C: Geneticks:  
A: Gene: sualaphap  
C: Superfamily: integrin alpha-2b chain  
C: Keywords: calcium binding; embryo; glycolysis; glycoprotein; heterodimer; transmembran

Query Match 11.6%; Score 719.5; DB 2; Length 1054;  
Best Local Similarity 23.1%; Pred. No. 1.4e-39;  
Matches 292; Conservative 178; Mismatches 428; Indels 365; Gaps 48;

QY 20 TDTFMDTRKRVIPGSRRTAFPGYTVQCHDSISGNWLVVVGAPLE-----TNGVQKTGD 72  
DB 19 TVAFNFDLRAPVKFDGPGQSLGFSVAQHRDQNTDWILGAPFAPTTQPGVTNG---GA 74

QY 73 VYKCPVI-----HGKCTKLNGLRVTLNSVSEKDNMRLGLSLATNPKDNSFLACSP--LWS 126  
DB 75 VYKCPVTLPSGSPCEQVFPDTTGNTEVLD-KSNQWFGATLASSGPDGRILACAPRLVWL 133

QY 127 HECGSSYTTGMCNRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSIYPMVEVQHFLI 186  
DB 134 Q-----TSTISPT----- 141

QY 187 NILKKEFYIGPGQIQGVWVQYGDVHVEFLNDYRSVKDVEAAASHIEORGTEETARGI 246  
DB 142 -----DKREPTGTCP-- 152

QY 247 EFARSEAFQKGRKGAKKVMIVITDGESHSDPLEKVIQOQSERDNVTRYAVAVLYNNR 306  
DB 153 ----- 152

QY 307 GINPETFLEIKYIASDPDDKHFNVTDEAALKDIDVALGDRIPLSEGTNKNETSFGLEM 366  
DB 153 -VGHSDFTNFVNSPCQSTDRDLF-----GFDK-----ITH 182

QY 367 SQTGSSSHVVDG--VLLGAVGAYDNGAVLAKETSAGKVIPLRESYLKEFPPE-LKNEGA 423  
DB 183 CQAGSAQIPSDNSTLWMCAPGSYVLOQIQAQSL-----TLDVSNTPQAVAFDPS 236

QY 424 YLGVTYTSVSSRQG-RVTVAGAPRPH-TGKVILFTMNNRSLT-IHQAMRGQOIGSYP 480  
DB 237 YRGYSIALGDFNGDGLDYVVGTPRGSRLGLVAIF---DQSLVEIITPVVVGQIVSYP 292

QY 481 GSEITSVDLDGQVTDVLLVAGPMYFN-----EGRERGKVVVYVELRQNRFFVYNGTKDQSH 535  
DB 293 GYSVASDVNGDGLDD-LLVGAPMFTNREPATEKKEAGRVYI--LQNA-----DH 340

QY 536 SYQ-----NARPGSSIASVRDLNQDNDVNDVVVGAPLE-DNHAGAIYIFHGFRGSI 584  
DB 341 SLGAPQMLTGKKIRARFGFPITSIGDSNQDGFNDVAIGAPYDGEDNSGVVYIYHGSARGL 400

QY 585 LKTPKQRIITASBLA-TGLQYFGCSIHQQLDNLNEDGLIDLAVCA--LGNRVILWSPVYQI 641  
DB 401 RUTESQVLTPSELGSDITTPGFSVDGQDMQNDVPDLVWGAESADRAILVTRTPVVL 460

QY 642 NASLHPEPSKINIFRDK-RSGRDATCLAAFLCFTPIPLAPHFQTTTGWIRYNATMDER 700  
DB 461 BAELTIEPIGINLNDKTYELPDGVTWTSFVAVACFI--VTGNHL-PARIGISYTLTVDS 517

QY 701 RYTPRAHL---DEGGRPTNRAVLLSSGOELCERINFVLDTP-ADYVKEPTFSVYSLED 756  
DB 518 ITSGRALLLEVLSQVTKQRN--LVNGMKPCDPLRAYVYNTIQDKLTPIAVDLQYELTD 575

QY 757 P-----DHGPMDDGWPPTLAVSPFMWGCNEDEHCVFDLVLDARSDLPTAMEYCORVL 810  
DB 576 ESILPEYILPINKAEVSSQTKVSIQNNC-VNNICIPIGITVTNLPN----- 625

QY 811 RKPQDCSAVTLGSPDFTTFFIESTRORVAVEATLENGENAYSTVLNISQSANLQFASLI 870  
DB 626 -----IVIGQAOELTLVVSINNEGDAPOSTLAVYYPEGLQYVRL 666

QY 871 QKEDSDGSIQVNEERLQOVNVSYPFRKAKAVAFRLD-SFSPKSIIFLHLEIELAA 929  
DB 667 RRANMDFSVTC-TEDSALRMITCDTGNPLV-KYTNLEFGLTLSTLQVSGDKDNIEFYLVA 724

QY 930 GSDSNERDSTKEDNVAPLRFHLKYEADVLFRSS--SLSHVEVKNLSLERYD----- 980  
DB 725 GSENNEDPTLNDNMLNVTAVIVDATLKLSSASYPEIVTYRVPEDNIVPEPTKNASEA 784

QY 981 GIGPPFSCIFRIQNLGLFPIHGMKMITIPIATRSNRLKLKRDPLTDEANTSCINWNS 1040  
DB 785 DIGMEVVLHYEVNRTGSSNAAEVTLNIRWPKDENGDFLYFLLGIMTDEGVT-CQI--SQ 841

QY 1041 TEYRPTVEBDLRAPOLNHSNDV-----VSINCNI- 1072

DB 842 QGANPLGVKLEASTKEQLSNSSTTOVSGRRKREGEYAEALAAQAEPICTPESCVLINCTID 901

QY 1073 RLVPNQBINFHLGNLWLSL-KALKYKSKMIWNAA-----LQRFHSPFIFREE 1122

DB 902 EIKATKSKVVRILGRFWERTFQKAVSEAVPIQVTLASTATATVRSIPYNIPLPMBPTDS 961

QY 1123 DPSQIEFEISKQEDWQVI-----WIIVSTLGLGLLALLVLALRKLGPFRSARRREP 1178

DB 962 TKASTL---ITABELVLPVSVIAMIWIIIVSVLGGIILLIILGLWKGFF-----ERKKP 1014

QY 1179 GLD 1181  
DB 1015 GED 1017

Search completed: June 24, 2004, 17:57:27  
Job time : 33 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OK protein - protein search, using sw model

Run on: June 24, 2004, 17:53:06 ; Search time 18 Seconds  
(without alignments)  
3436.630 Million cell updates/sec

Title: US-09-980-403-2  
Perfect score: 6224  
Sequence: 1 MDLPRLVAVALSILWPGT.....FRSARRRREPLDTPPKVLE 1188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141631 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 6188.5 | 99.4        | 1189   | 1     | ITAH_HUMAN  |
| 2          | 2403.5 | 38.6        | 1167   | 1     | ITAG_HUMAN  |
| 3          | 2159   | 34.7        | 1151   | 1     | ITAI_HUMAN  |
| 4          | 2149.5 | 34.5        | 1180   | 1     | ITAI_RAT    |
| 5          | 1910.5 | 30.7        | 1178   | 1     | ITAI_MOUSE  |
| 6          | 1901.5 | 30.6        | 1170   | 1     | ITAI_BOVIN  |
| 7          | 1863   | 29.9        | 1181   | 1     | ITAI_HUMAN  |
| 8          | 1146.5 | 18.4        | 1170   | 1     | ITAI_HUMAN  |
| 9          | 1139   | 18.2        | 1162   | 1     | ITAI_HUMAN  |
| 10         | 1132.5 | 18.2        | 1163   | 1     | ITAI_MOUSE  |
| 11         | 1104   | 17.7        | 1153   | 1     | ITAI_HUMAN  |
| 12         | 1093.5 | 17.6        | 1152   | 1     | ITAI_MOUSE  |
| 13         | 1073.5 | 17.2        | 1163   | 1     | ITAI_MOUSE  |
| 14         | 994    | 15.3        | 1167   | 1     | ITAI_MOUSE  |
| 15         | 924.5  | 14.9        | 1179   | 1     | ITAI_HUMAN  |
| 16         | 825    | 13.3        | 285    | 1     | ITAI_CHICK  |
| 17         | 799.5  | 12.8        | 1035   | 1     | ITAI_HUMAN  |
| 18         | 693.5  | 11.1        | 1039   | 1     | ITAI_MOUSE  |
| 19         | 673.5  | 10.8        | 1032   | 1     | ITAI_XENLA  |
| 20         | 672.5  | 10.8        | 1038   | 1     | ITAI_HUMAN  |
| 21         | 611.5  | 9.8         | 1130   | 1     | ITAI_HUMAN  |
| 22         | 607.5  | 9.8         | 1034   | 1     | ITAI_CHICK  |
| 23         | 606    | 9.7         | 1049   | 1     | ITAI_HUMAN  |
| 24         | 604.5  | 9.7         | 1053   | 1     | ITAI_MOUSE  |
| 25         | 597.5  | 9.6         | 1066   | 1     | ITAI_HUMAN  |
| 26         | 594    | 9.5         | 1053   | 1     | ITAI_MOUSE  |
| 27         | 584.5  | 9.4         | 1044   | 1     | ITAI_MOUSE  |
| 28         | 580.5  | 9.3         | 1066   | 1     | ITAI_CRISP  |
| 29         | 579    | 9.3         | 1091   | 1     | ITAI_MOUSE  |
| 30         | 577.5  | 9.3         | 1050   | 1     | ITAI_XENLA  |
| 31         | 568    | 9.1         | 1072   | 1     | ITAI_CHICK  |
| 32         | 566.5  | 9.1         | 1146   | 1     | ITAI_DROME  |
| 33         | 563    | 9.0         | 1044   | 1     | ITAI_CHICK  |

|    |       |     |      |   |            |
|----|-------|-----|------|---|------------|
| 34 | 549.5 | 8.8 | 1000 | 1 | ITAS_DROME |
| 35 | 546.5 | 8.8 | 1226 | 1 | PAT2_CABEL |
| 36 | 542   | 8.7 | 1115 | 1 | ITAT_DROME |
| 37 | 537   | 8.6 | 1048 | 1 | ITAV_HUMAN |
| 38 | 525.5 | 8.4 | 1181 | 1 | ITAT_HUMAN |
| 39 | 520.5 | 8.4 | 1179 | 1 | ITAT_MOUSE |
| 40 | 515.5 | 8.3 | 1139 | 1 | ITAI_CABEL |
| 41 | 509.5 | 8.2 | 1025 | 1 | ITAT_HUMAN |
| 42 | 502.5 | 8.1 | 1106 | 1 | ITAT_RAT   |
| 43 | 487   | 7.8 | 1033 | 1 | ITAT_MOUSE |
| 44 | 485.5 | 7.8 | 1015 | 1 | ITAT_DROME |
| 45 | 472.5 | 7.6 | 1039 | 1 | ITAT_HUMAN |

ALIGNMENTS

RESULT 1  
ITAH\_HUMAN  
ID ITAH\_HUMAN STANDARD; PRT; 1189 AA.  
AC Q9UKX5; Q9UKQ1;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-11 precursor.  
GN ITGAl1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal heart, and Osteoblast;  
RX MEDLINE=99417678; PubMed=10486209;  
RA Lehnert K., Ni J., Leung E., Gough S.M., Yao W.P., Liu D.,  
Wang S.-X., Morris C.M., Krissansen G.W.;  
RT "Cloning, sequence analysis, and chromosomal localization of the novel  
human integrin alpha11 subunit (ITGAl1).";  
RL Genomics 60:179-187(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal muscle, and Uterus;  
RX MEDLINE=99395147; PubMed=10464311;  
RA Velling T., Kusche-Gullberg M., Sejersen T., Gullberg D.;  
RT "cDNA Cloning and Chromosomal Localization of Human alpha(11)  
Integrin. A collagen-binding, i domain-containing, beta(1)-associated  
integrin alpha-chain present in muscle tissues.";  
RL J. Biol. Chem. 274:25735-25742(1999).  
RN [3]  
RP SEQUENCE OF 954-1189 FROM N.A.  
RC TISSUE=Fibroblast;  
RA Andreu N., Estivill X., Escarceller M., Sumoy L.;  
RX Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11  
CC ASSOCIATES WITH BETA-1.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND  
CC HEART. INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO  
CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO  
CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,  
CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN  
CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.  
CC -!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING  
CC FETAL MUSCLE CELLS (IN VITRO).  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSEPT) IS A VMFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VMFA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL; AF109681; AAP01258.1; -  
CC EMBL; AF137378; AAD51919.2; -  
CC EMBL; AL353064; CAB94392.1; -  
CC HSSP; P17301; IAOX.  
CC -----  
CC DR MIM; 604789; -  
CC DR GO; 0008305; C: integrin complex; TAS.  
CC DR GO; 0004895; F: cell adhesion receptor activity; TAS.  
CC DR GO; 0005518; F: collagen binding; TAS.  
CC DR GO; 0007160; P: cell-matrix adhesion; TAS.  
CC DR GO; 0007517; P: muscle development; TAS.  
CC DR InterPro; IPR000413; Integrin\_alpha.  
CC DR InterPro; IPR002035; VWF A.  
CC DR Pfam; PF01839; FG-GAP; 3.  
CC DR Pfam; PF00092; vwa; 1.  
CC DR PRINTS; PR01185; INTEGRINA.  
CC DR PRINTS; PR00453; VWFADOMAIN.  
CC DR SMART; SM00191; Int\_alpha; 5.  
CC DR SMART; SM00327; VWA; 1.  
CC DR PROSITE; PS00242; INTEGRIN\_ALPHA; FALSE\_NEG.  
CC DR PROSITE; PS0234; VWA; 1.  
CC DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
CC Signal; Repeat; Calcium; Magnesium; Polymorphism.  
CC SIGNAL; 1 22  
CC CHAIN 23 1189  
CC DOMAIN 23 1142 INTEGRIN\_ALPHA-11.  
CC TRANSMEM 1143 1165 EXTRACELLULAR (POTENTIAL).  
CC DOMAIN 1166 1189 CYTOPLASMIC (POTENTIAL).  
CC REPEAT 38 94 FG-GAP 1.  
CC REPEAT 102 163 FG-GAP 2.  
CC DOMAIN 167 345 VWA.  
CC REPEAT 359 420 FG-GAP 3.  
CC REPEAT 422 475 FG-GAP 4.  
CC REPEAT 477 537 FG-GAP 5.  
CC REPEAT 539 598 FG-GAP 6.  
CC REPEAT 601 653 FG-GAP 7.  
CC DOMAIN 1154 1162 POLY-LEU.  
CC DOMAIN 1174 1177 POLY-ARG.  
CC CA\_BIND 488 496 POTENTIAL.  
CC CA\_BIND 551 559 POTENTIAL.  
CC DISULFID 76 83 BY SIMILARITY.  
CC DISULFID 121 139 POTENTIAL.  
CC DISULFID 129 159 POTENTIAL.  
CC DISULFID 659 688 BY SIMILARITY.  
CC DISULFID 674 723 BY SIMILARITY.  
CC DISULFID 781 787 BY SIMILARITY.  
CC DISULFID 881 893 BY SIMILARITY.  
CC CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 358 358 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 449 449 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 462 462 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 528 528 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 642 642 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 694 694 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 857 857 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 884 894 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 973 973 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 1032 1032 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 1040 1040 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC V -> M.  
CC VARIANT 433 433 /FTid=VAR\_009889.  
CC VARIANT 524 524 R -> L.  
CC VARIANT /FTid=VAR\_009890.

FT VARIAT 972 972 L -> P.  
FT VARIAT 1003 1003 /FTid=VAR\_009891.  
FT VARIAT 1030 1030 I -> M.  
FT VARIAT 1030 1030 /FTid=VAR\_009892.  
FT VARIAT 1030 1030 Missing.  
FT VARIAT 1094 1094 /FTid=VAR\_009893.  
FT VARIAT 1094 1094 L -> V.  
FT VARIAT 1094 1094 /FTid=VAR\_009894.  
SQ SEQUENCE 1189 AA; 133609 MW; 50303C08A44CD52 CRC64;  
  
Query Match  
Best Local Similarity 99.4%; Score 6188.5; DB 1; Length 1189;  
Matches 1184; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
  
QY 1 MDLPRGLVVAWALSILWPGFTTFNMDTRKPRVIPSRTAFPGYTVQOHDISGNKWLVA 60  
DB 1 MDLPRGLVVAWALSILWPGFTTFNMDTRKPRVIPSRTAFPGYTVQOHDISGNKWLVA 60  
QY 61 PLENGYOKTGDVYKCPVIHGNCTKLNLRVTLNVSERKDNMRLGLSLATNPKDNSPLA 120  
DB 61 PLENGYOKTGDVYKCPVIHGNCTKLNLRVTLNVSERKDNMRLGLSLATNPKDNSPLA 120  
QY 121 CSPLWSHECGSSYTTGKCSRVSNSFRPSKTVAPALORCOTYMDIVIVLDSNSIYPWE 180  
DB 121 CSPLWSHECGSSYTTGKCSRVSNSFRPSKTVAPALORCOTYMDIVIVLDSNSIYPWE 180  
QY 181 VOHFLINILKFKYIGPGQIOGVVQYGEDVVEHFLNDYRSVKDVEAASHIEORGGTET 240  
DB 181 VOHFLINILKFKYIGPGQIOGVVQYGEDVVEHFLNDYRSVKDVEAASHIEORGGTET 240  
QY 241 RTAFGIEFARSEAPQKGRGAKKVMIVITDGSHTSPDLKVIQOSERDNVTRYAVAVL 300  
DB 241 RTAFGIEFARSEAPQKGRGAKKVMIVITDGSHTSPDLKVIQOSERDNVTRYAVAVL 300  
QY 301 GYNRRGINPETFLNEIKYIASDDPKHFFNVTDAAALKDIVDALGDRIFSLGNTKNET 360  
DB 301 GYNRRGINPETFLNEIKYIASDDPKHFFNVTDAAALKDIVDALGDRIFSLGNTKNET 360  
QY 361 SFGLEMSQTGFSSHVDEGVLLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEPPELKN 420  
DB 361 SFGLEMSQTGFSSHVDEGVLLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEPPELKN 420  
QY 421 HGAVLGYVTTSVSSROGRVYVACAPRNTGKVLFTMNNRSLTTHOAMRGQIQISYF 480  
DB 421 HGAVLGYVTTSVSSROGRVYVACAPRNTGKVLFTMNNRSLTTHOAMRGQIQISYF 480  
QY 481 GSEITSDIDGDGVTDLVNGAPMYFNEGRERGKVVYVELRQNEFVYNGTLKDSHSYQNA 540  
DB 481 GSEITSDIDGDGVTDLVNGAPMYFNEGRERGKVVYVELRQNEFVYNGTLKDSHSYQNA 540  
QY 541 RFSSIASVRDLNODSYNDVVGAPLEDNEAGALYIFHGFRGSILKTPKQBITASBLATG 600  
DB 541 RFSSIASVRDLNODSYNDVVGAPLEDNEAGALYIFHGFRGSILKTPKQBITASBLATG 600  
QY 601 LQYFGCSIHGOLDLNEDELIDLVAGLGNVILWSRPVQINASLHFPSPKINIFHRDCK 660  
DB 601 LQYFGCSIHGOLDLNEDELIDLVAGLGNVILWSRPVQINASLHFPSPKINIFHRDCK 660  
QY 661 RSGDATCLAAFLCFTPIFLAPHQTTVGIRYNATWDERRYTPRAHLDEGGDPTTRAV 720  
DB 661 RSGDATCLAAFLCFTPIFLAPHQTTVGIRYNATWDERRYTPRAHLDEGGDPTTRAV 720  
QY 721 LLSSGQELCBRINPHVLDADYVKPVTFSVEYSLEDPDHGMPLDDGWPTTLRVSPVFWNG 780  
DB 721 LLSSGQELCBRINPHVLDADYVKPVTFSVEYSLEDPDHGMPLDDGWPTTLRVSPVFWNG 780  
QY 781 CNEDEHCVPLVLDARSPLTAMEYCORVLRKPAQDCSAVTLSPDTTVTFIESTRQRAV 840  
DB 781 CNEDEHCVPLVLDARSPLTAMEYCORVLRKPAQDCSAVTLSPDTTVTFIESTRQRAV 840  
QY 841 EATLENGENAYSTVLNISQSANLQPASLIQKSDSGSIECVNEERELQKQCNVSPFF 900  
DB 841 EATLENGENAYSTVLNISQSANLQPASLIQKSDSGSIECVNEERELQKQCNVSPFF 900



QY 901 RAKAKVAFRLDSEFSKIFLHLEIELELAAGSDSNRSTKEDNVAPLRPHLKYEADVLF 960  
 DB 902 RAKAKVAFRLDSEFSKIFLHLEIELELAAGSDSNRSTKEDNVAPLRPHLKYEADVLF 960  
 QY 961 RSSLISHVEKLNLSLVDGIPGPPSCIFRIQNLGLPIHGMKKITIPATSGNRL 1020  
 DB 961 RSSLISHVEKLNLSLVDGIPGPPSCIFRIQNLGLPIHGMKKITIPATSGNRL 1020  
 QY 1021 KLRLDTDE-ANTSNCINMGSTYRPTPVVEEDLRAPQLNHSNDSVWVINCINIRLVPNQ 1079  
 DB 1021 KLRLDTDEVANTSCINMGSTYRPTPVVEEDLRAPQLNHSNDSVWVINCINIRLVPNQ 1080  
 QY 1080 INFHLGNLWLSLXALKYKSKIMVNAALQROPHSPFIFREDEPSRQIBFISKQEDWQ 1139  
 DB 1081 INFHLGNLWLSLXALKYKSKIMVNAALQROPHSPFIFREDEPSRQIBFISKQEDWQ 1140  
 QY 1140 VPIWIVGSTGLGLLLALLVLALRKLGFPSARRRREPGDPTPKVLE 1188  
 DB 1141 VPIWIVGSTGLGLLLALLVLALRKLGFPSARRRREPGDPTPKVLE 1189

RESULT 2

ITAG HUMAN STANDARD; PRT; 1167 AA.  
 AC 075578; 09HE28;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Integrin alpha-10 precursor.  
 GN ITGA10.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Articular chondrocytes;  
 RX MEDLINE=98352078; PubMed=9685391;  
 RA Camper L., Hellman U., Lundgren-Aakerlund E.;  
 RT "Isolation, cloning, and sequence analysis of the integrin subunit  
 RT alpha10, a beta1-associated collagen binding integrin expressed on  
 RT chondrocytes.";  
 RL J. Biol. Chem. 273:20383-20389 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Endothelial cells, and Heart;  
 RX MEDLINE=20169197; PubMed=10702680;  
 RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,  
 RA Wang S.-X., Langley R., Krissansen G.W.;  
 RT "The integrin alpha10 subunit: expression pattern, partial gene  
 RT structure, and chromosomal localization.";  
 RL Cytogenet. Cell Genet. 87:238-244 (1999).  
 CC -!- FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.  
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10  
 CC ASSOCIATES WITH BETA-1.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: Widely expressed with highest expression in  
 CC muscle and heart. Found in articular cartilage.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMPA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VMPA domain.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AF074015; AAC31952.1; -.  
 DR EMBL; AF112345; AAF21944.1; -.  
 DR EMBL; AF172723; AAF61638.1; -.  
 DR HSSP; P17301; IAOX.  
 DR Genew; HGNC:6135; ITGA10.  
 DR MIM; 604042; -.  
 DR GO; GO:0008305; C: integrin complex; TAS.  
 DR GO; GO:0004895; F: cell adhesion receptor activity; TAS.  
 DR GO; GO:0005518; F: collagen binding; TAS.  
 DR GO; GO:0007160; P: cell-matrix adhesion; TAS.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWF\_A.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00092; vwa; 1.  
 DR PRINTS; PRO1185; INTEGRINA.  
 DR PRINTS; PRO4453; VWFADOMAIN.  
 DR SMART; SM00191; Int\_alpha; 4.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; FALSE\_NEG.  
 DR PROSITE; PS0234; VMPA; 1.  
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Signal; Repeat; Calcium; Magnesium.  
 FT SIGNAL 1 22  
 FT CHAIN 23 1167  
 FT DOMAIN 23 1122  
 FT TRANSMEM 1123 1145  
 FT DOMAIN 1146 1167  
 FT REPEAT 38 97  
 FT REPEAT ?  
 FT DOMAIN 167 350  
 FT REPEAT 365 427  
 FT REPEAT 428 482  
 FT REPEAT 483 545  
 FT REPEAT 546 605  
 FT REPEAT 608 660  
 FT DOMAIN 1134 1140  
 FT CA\_BIND 494 502  
 FT CA\_BIND 558 566  
 FT CA\_BIND 620 628  
 FT DISULFID 76 86  
 FT DISULFID 666 675  
 FT DISULFID 681 736  
 FT DISULFID 789 795  
 FT CARBOHYD 98 98  
 FT CARBOHYD 234 234  
 FT CARBOHYD 336 336  
 FT CARBOHYD 364 364  
 FT CARBOHYD 733 733  
 FT CARBOHYD 763 763  
 FT CARBOHYD 839 839  
 FT CARBOHYD 921 921  
 FT CARBOHYD 1011 1011  
 FT CARBOHYD 1018 1018  
 FT CARBOHYD 1039 1039  
 FT CARBOHYD 844 844  
 FT CONFLICT 909 909  
 FT CONFLICT 926 926  
 SQ SEQUENCE 1167 AA; 127573 MW; A87D3A1C25C1AE0 CRC64;  
 Query Match 38.6%; Score 2403.5; DB 1; Length 1167;  
 Best Local Similarity 43.0%; Pred. No. 6.4e-148;  
 Matches 514; Conservative 212; Mismatches 425; Indels 45; Gaps 16;  
 QY 1 MDLPRGLVAVALSLSWPGFTDTFMMDTRKPRVPGSTAPFGYTVQGHDSGNLWVGA 60  
 DB 1 MELPFVTHFLPLVFLTGLCSPPNLDEHPLRFPPEAEFGYSLQHVGGQRMVLVGA 60  
 QY 61 PLENTGYQKTDVYKCPV--IHGNCFKLNLGRVTLNSVSKDNMELGLSLATNPKDNS 117  
 DB 61 PWDGSPDRGDVYRCVPGGANAPCAKHLGDYQLGNSSHPAVMHLGKSLLETGDDG 120  
 QY 118 FLACSPWLSHEGGSYTTGMCGRVNSNFRFSKTVAPALQRCQYTMIVIVLDSNSIYP 177



Db 121 FMACFLWSRACGSSVSSGICARVDASFQOGLSLAPTAQRCPTMYDVVIVLDGNSIYP 180  
 Qy 178 KYEVOHFLNLKFFVIGPQIQGVVGVGEVDWHEFLHNDVRSKVDVVEAAASHIEORG 237  
 Db 181 WSEVQTFRLVKCLFIDPQIQVGLVGVGEVPHWLSGDFRTKEVVRKAANLSRREG 240  
 Qy 238 TETRTAFGTFEFAKSEAFQK--GGRKAKKAMVIVITDGESHSDPLEKVIQOQRDNVTRY 295  
 Db 241 RETKTAQAMVACTEGFSQSHGGRPPAARLLVVVTDGESHDEGELPAALKACAGRVTRY 300  
 Qy 296 AVAVILGYNNRGINPTFTLEIKVIASDDPDGKFFPNVDEAALKDIDVDAIGDRIFSLEGT 355  
 Db 301 GIAVLGHLRQRDPSPFLREHTIASDDPDERFFNVDEAALTIDVDAIGDRIFGLEGS 360  
 Qy 356 N-KNETSPGLEMSQTFSSHVDGVLGAVGAYDNGAVLKETSAGKVIPIRESVLKEF 414  
 Db 361 HAENESSFGLEMSQIGFSTHRLKDGILFGMVGAYDNGGVLWLEGHRLFFPRMALEDEF 420  
 Qy 415 PEELKXHGAVLGVTYSVSSRGRVYVAGAPRENTGKVLFTHNNESLTIHQAMRQ 474  
 Db 421 PPAQNHAAVLGVSVSSMLRGGRLFLSGAPRFRGRGVIAFOLKKGAGRVAVASLQGE 480  
 Qy 475 QIGSYFGSEITSVIDIDGDGVTDLVLYGAPMYFN--EGRERGVYVYEL--RQNRVYNGTLK 532  
 Db 481 QIGSYFGSELCPDLTDRTDGTDLVLAAPMFLGPQNKETGRVYVLYVGGQSLTLTQGTIQ 540  
 Qy 533 DSHSYQVNRGSSIASVRLNODSYNDVYVVGAPLEDNHAGAIYIPHGFGSILKTPKORI 592  
 Db 541 PEPF--QDARPGFAGPALPDNLQGFADVAVGAPLEDHGOALYLYHGTSQVGRHPAQRI 599  
 Qy 593 TABELATGLQYFGCSIHQGLDINEDGLIDLAVALGNNAVILMSRPVVOQINASLHFPSPKI 652  
 Db 600 AASMPHNLVSFGSVDRGLDLOGDLDLVAVGVAQGAAILLSRPIVHLTSPLEVTFOAI 659  
 Qy 653 NIFRDCRGRDATCLAPLCTPIFLAPHFTTTVIGIRYNAITDERRYTRAHLDEGG 712  
 Db 660 SVVQRDCRRGQBAVCLTAALCFQVTSRTIPGRWDHQFYMRFTASLDEWTAGARAAPDGG 719  
 Qy 713 DRTNRAVLSSGOELCERINFVLDTADVYKVPFSEVYSLEDDPH--GPMLDGDGPTTL 771  
 Db 720 QRLSPRLRLSVGNVTCQQLHFVLDTSVLRVALVTVPALDNTTKRGPVLNKGSPSTSI 779  
 Qy 772 RVSVFVNGCNEHCVDPDLVLAARSDLTAMBYCORVLKPAQDCSAYTLSPDTTFFII 831  
 Db 780 QKLVFSPKDCGPDNECVTLVLQVNMIDI-----RGRK-----APFVV 817  
 Qy 832 ESTQRVAVENTLENGENAVSTVNIQSANLOFASLICKEDSDGSIKWCNEERLQX 891  
 Db 818 RGRKRVLTSTLENKXENAYNTSLIIFSRNLHSLITPQRESPIKVECAAPS--AHAR 875  
 Qy 892 VCNVSYFFRAKAVFRLDSEFSKSIPLHLEIELAAGSDSNEDSTKEDNVAPLRFHL 951  
 Db 876 LCSVGHVPVOTGAKVTELLFEFSCSLLSQVFGKLTASSDSLSERNGLTQNTAQTSAVI 935  
 Qy 952 KYEADVLFTRSSLSHVEKLNLSERYDGIQPPFCIFRIONLGLPPIHGMKTIPI 1011  
 Db 936 QYEPHLLFSBESTLHYREHPYGLP--VVGPEFKTTLRVQNLGCVVSGLLISALLPA 993  
 Qy 1012 ATRSGNLLKRLDPLTDEANTSCNIMGNSTYEPTPVE--EDLRRAPQLNHSNDVVSINC 1070  
 Db 994 VAHGNVFLSLQVITN--NASC--IVQNLTEPPGPPVHPEELQHTNRLNGSNQCCQVVR 1050  
 Qy 1071 NI--RLVFNQENFILLGNLWLSLKALKYKSMKWNAALQRFQHSFFIREDDPSQIE 1129  
 Db 1051 HLGQAKGTEVSVGLLRLVHNEFFRAKFSLTVTVSTFELGTGEGSVLQTEASRWSES 1110  
 Qy 1130 FEISKQEDWQVPIWIVGSTIGGLLALLVLAALRKLGFRRSAR----RRREGLD 1181  
 Db 1111 LEVQTPRILISLWILGSLVGLLALLVFLCLWKLGFPAHKKIPEEXREKLE 1166

RESULT 3

ITAL\_HUMAN

ITAL\_HUMAN STANDARD; PRT; 1151 AA.  
 P56199;  
 01-NOV-1997 (Rel. 35, Created)  
 01-NOV-1997 (Rel. 35, Last sequence update)  
 10-OCT-2003 (Rel. 42, Last annotation update)  
 Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).  
 GN ITGAL.  
 Homo sapiens (Human).  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 [1]  
 SEQUENCE FROM N.A. PubMed=8428973;  
 MEDLINE=93155124; Epstein M.R.; Marcantonio E.E.;  
 Brissewitz R.; Expression of native and truncated forms of the human integrin alpha  
 1 subunit.";  
 J. Biol. Chem. 268:2989-2996(1993).  
 CC !- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND  
 COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-  
 E-R IN COLLAGEN  
 CC !- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1  
 ASSOCIATES WITH BETA-1.  
 CC !- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC !- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC !- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC !- SIMILARITY: Contains 1 VWFA domain.  
 CC !- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC !- DATABASE: NAME=PROB; NOTE=CD guide CD49a entry;  
 WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49a.htm".  
 PIR; A45226; A45226.  
 PDB; 1OC5; 17-MAY-00.  
 Genew; HGNC:6134; ITGAL.  
 MIM; 192968; --  
 GO; GO:0008305; C:integrin complex; TAS.  
 GO; GO:0004895; F:cell adhesion receptor activity; NAS.  
 GO; GO:0005518; F:collagen binding; TAS.  
 GO; GO:0007160; P:cell-matrix adhesion; NAS.  
 InterPro; IPR000413; Integrin\_alpha.  
 InterPro; IPR002035; VWF\_A.  
 Pfam; PF01839; FG-GAP; 3.  
 Pfam; PF00357; integrin\_A; 1.  
 Pfam; PF00092; vwa; 1.  
 SMART; SM00191; int\_alpha; 5.  
 SMART; SM00327; VWA; 1.  
 PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 PROSITE; PS0234; VWFA; 1.  
 Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 Repeat; Calcium; Magnesium; 3D-structure.  
 Repeat; Extracellular (POTENTIAL).  
 DOMAIN 1 1113  
 FT TRANSMEM 1114 1136  
 FT DOMAIN 1137 1151  
 FT REPEAT 16 75  
 FT REPEAT ? ?  
 FT REPEAT ? ?  
 FT REPEAT 147 360  
 FT REPEAT 349 404  
 FT REPEAT 405 427  
 FT REPEAT 459 520  
 FT REPEAT 540 599  
 FT REPEAT 602 654  
 FT REPEAT 470 478  
 FT CA\_BIND 552 560  
 FT CA\_BIND 614 622  
 FT SITE 1139 1142  
 FT DISULFID 54 64  
 FT DISULFID 660 669  
 FT DISULFID 675 728  
 FT DISULFID 780 786  
 FT DISULFID 850 858  
 FT DISULFID 1002 1034  
 FT DISULFID 1037 1044  
 FT CARBOHYD 46 46  
 N-LINKED (GLCNAC... ) (POTENTIAL).

FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 504 504 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 720 720 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 752 752 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 855 855 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 880 880 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 887 887 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 911 911 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 938 938 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 946 946 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 980 980 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 SQ SEQUENCE 1151 AA; 127837 MW; 6B3F3C1AABF52808 CRC64;

Query Match 34.7%; Score 2159; DB 1; Length 1151;  
 Best Local Similarity 38.3%; Pred. No. 4.7e-132;  
 Matches 457; Conservative 235; Mismatches 417; Indels 84; Gaps 20;

QY 23 FMMDTRPRVPGSRAPFGYTTVOQHDISGNKWLTVGAPLETNGYQKTDGVYKCPVHGN 82  
 DB 1 FNVVDKSNMTSGPVEDVFGYTVQYENBEKWLIGSPVGVQPNRTGVDYKCPVGRGE 60

QY 83 ---CTKILNG-RVLSNVSEKONRGLGLSLATNPKNLSFLACPLSHGEGSSYTTGM 138  
 DB 61 SLPCVKLDELNTSIPNTEYKENVTFGSTLVTPN-NGGFLACGLPYAYRCGHLHYTTGI 119

QY 139 CSRVNSNFRSKTAPALQRCQTYMDIVILVDGNSIYPPWVEVQHFLINILKKEVIGPGQ 198  
 DB 120 CSDVSPPTQVNSIAP-VQECSTQLDIVILVDGNSIYPPWDSVTAFLNDLLKRWMDIGPKQ 178

QY 199 IQGVVQVGEDVWHEFHNDTRSKVDVVEASHIIEQRCGTTRTAFGIEFARSAP--QK 256  
 DB 179 TQGVGVQVGENWHEFNILKYSSTEEVLVAKKIVQRCGRQTMALGTDTRAKFAFTAR 238

QY 257 GGRKGAKKMIVITDGBSHDSDLEKVIQOQSERDNTVYAVVLGYNRRGINPETPLNE 316  
 DB 239 GARRGVKKVMIVTDGBSHDHRLLKVIQDCDENIQRFSAIIGSYNRRGLSTKTFVBE 298

QY 317 IKYIASDDPKHFNVTDEAALKDVIDALGRIPFLEGT-NKNETSFGLEMSQTFSSHV 375  
 DB 299 IKSIASEPTKHFNVSDDELAVTIVTKTLGERIFALREATAQCSAASPEMEXSQTFSAHY 358

QY 376 VEDGVILCAVCAVNGAVLAKETSAGKVIPLRESVLKPEPPELKNHGAIVGYTVSVSS 435  
 DB 359 SODWMLCAVCAVNGVTVVQKASQIIIPNTTFNVESTKKNPLASLYLGYTVSNATAS 418

QY 436 RQGRVYVAGAPRPHGTGKVLFTMHNRSLSLTHQARQQOIGSYFGSISVTDIGDGV 495  
 DB 419 SGDVLVIAGQPRYNTGQVILYRVEDG-NIKILQTLGSGEIGSYFGSILITTDIDKDSNT 477

QY 496 DVLIVGAPMYP-NEGREGKRYVVELRQNRVYNGTLK-----DSHYQN--- 539  
 DB 478 DILAVGAPMYPGTEKEEGKRYVVALNQTRFEYQMSLEPIKQTCSSRQHNCSCTENKNE 537

QY 540 ---ARFGSSIASVDLNDSDVNDVVGAPLBDNHAGAIYIPHGRGSLIKTKPKORITASE 596  
 DB 538 PCGARFGTAAVAKDLNDGNDVIGAPLEDHGGAVIYHGGKTKIRKYAQRIPSGG 597

QY 597 LATGLQVPGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQVINALHPSPSKINIFH 656  
 DB 597 LATGLQVPGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQVINALHPSPSKINIFH 656

DB 598 DGKTLKPFQSGIHGEMDLNGDLTDVTIGGLGGAALFWSRDVAVVVKVTNFPNPKVNIQK 657  
 QY 657 RDKRSGRDATCLAAFLCFTPIFIAPHFQTTTIGIRYNATMDERRYTPRAHLDEGDRFT 716  
 DB 658 KNCHMEGKETVCINATVCFEVLKSKEDTIYEAQLQVRVTLDSLRQISRSFFSGTOBRKV 717  
 QY 717 NRALLSSGOELCERINFHVLDTADYKVPYTFPSVEYSLDDPHGMLDDGWPYTLRVSVP 776  
 DB 718 ORNITVARKSB--CTKGSFYMLDKHDFQDSVRITLDFNLTDPENGFVLDDSLPNSVHYIP 775  
 QY 777 FWNGCNDEHCVPDLVLDARSDLPTAMEYQORVLKPAQDCSAYTLTDFDTTTFVIESTRQ 836  
 DB 776 FAKDOGNKKEKISDL-----SLHVAITTEKDLLIVRSQND 809  
 QY 837 RVAYEATLENRGENAYSTVNLISOSANLQFASL--IOKEDSDGSGIECVNEERLQKQVCN 894  
 DB 810 KFNYSLVTKTKOSAYNTRTIVHSPNLFSGTGAIOKSCSN-----HNITCK 859  
 QY 895 VSYEPFFRAKAVAPRLDSBFSKSFHLHLELAAGSDNSNERDSTKDNVAPLFLHLYKE 954  
 DB 860 VGYPFLARGEMVTKILFQNTSVLMENVILYLSATSDSEPPETLSDNVNVIPIVYKE 919  
 QY 955 ADVLFTRSSSLSHVEVKLNS---LERYDIGIPFSCIPRIQNLGLPPIHGMKMTIP 1010  
 DB 920 VGLQFYSSAGEYHISIAANETVPEVINSTEDIGNEINIFYLIRKSGSGFPMPDELKLSIFP 979  
 QY 1011 IATFSGNRLKLRLDFTLEANTSC-----NINGNSTEYRPTVVEEDLRRAPQLNHSND 1064  
 DB 980 NMTSGYFVL-YPTGLSSSEANCRPHIFPDPSINSNGKMTSTDLKRGITLDCNTCK 1038  
 QY 1065 VWSINCINRLVPOQINFLHLLGNLWLSLKALKYKSMKIMVNAALORQHFSPFIREBDP 1124  
 DB 1039 PATITCNLTSSDISQVNVSLI--LWKPTFKSYVPSLSNLATIRGEL-RSENASVLSSNQ 1095  
 QY 1125 SRQIEPEISKQB-DWQVPIHIVGSTLGGILLALVLAALKGLGFPFSSARRR 1176  
 DB 1096 KRELAIQISKDGLGPRVPLWVLLSAPAGLULLLALWKIGFRRPLKCK 1148

RESULT 4  
 ITAL RAT STANDARD; PRT; 1180 AA.  
 AC P18614;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1)  
 DE (CD49a).  
 GN ITGA1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90338125; PubMed=2380249;  
 RA Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A.,  
 RA Eech F., Carboneito S., Reichardt L.F.;  
 RT "Molecular cloning of the rat integrin alpha 1-subunit: a receptor  
 RT for laminin and collagen."  
 RL J. Cell Biol. 111:709-720(1990).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.  
 RX MEDLINE=99313197; PubMed=1038626;  
 RA Nolte M., Pepinsky R.B., Vanyaminov S.Y., Koteliarsky V.,  
 RA Gotwals P.J., Karpusas M.;  
 RT "Crystal structure of the alpha1beta1 integrin I-domain: insights into  
 RT integrin I-domain function."  
 RL FEBS Lett. 452:379-385(1999).  
 CC -/- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND  
 CC COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-G-  
 CC E-R IN COLLAGEN.  
 CC -/- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1

ASSOCIATES WITH BETA-1.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -1- SIMILARITY: Contains 1 VWFA domain.  
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X52140; CAA36384.1; -.  
 DR PIR; A35854; A35854.  
 DR PDB; 1CK4; 03-MAY-00.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWFA.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00357; integrin\_A; 1.  
 DR Pfam; PF00092; vwa; 1.  
 DR SMART; SM00327; VWFA; 1.  
 DR SMART; SM00327; VWFA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS00234; VWFA; 1.  
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Signal; Repeat; Calcium; Magnesium; 3D-structure.  
 FT SIGNAL 29  
 FT CHAIN 29 1180  
 FT DOMAIN 29 1142  
 FT TRANSMEM 1143 1165  
 FT DOMAIN 1166 1180  
 FT REPEAT 44 103  
 FT REPEAT 44 103  
 FT REPEAT 175 388  
 FT REPEAT 377 432  
 FT REPEAT 433 484  
 FT REPEAT 485 565  
 FT REPEAT 567 626  
 FT REPEAT 629 681  
 FT CA\_BIND 497 505  
 FT CA\_BIND 579 587  
 FT CA\_BIND 641 649  
 FT SITE 1168 1172  
 FT SITE 82 92  
 FT DISULFID 687 696  
 FT DISULFID 702 755  
 FT DISULFID 807 813  
 FT DISULFID 877 885  
 FT DISULFID 1029 1062  
 FT DISULFID 1066 1073  
 FT CARBOHYD 100 100  
 FT CARBOHYD 105 105  
 FT CARBOHYD 112 112  
 FT CARBOHYD 217 217  
 FT CARBOHYD 317 317  
 FT CARBOHYD 341 341  
 FT CARBOHYD 402 402  
 FT CARBOHYD 418 418  
 FT CARBOHYD 459 459  
 FT CARBOHYD 531 531  
 FT CARBOHYD 698 698  
 FT CARBOHYD 747 747  
 FT CARBOHYD 779 779  
 FT CARBOHYD 820 820  
 FT CARBOHYD 839 839  
 FT CARBOHYD 882 882  
 FT CARBOHYD 907 907  
 FT CARBOHYD 938 938  
 FT CARBOHYD 965 965

FT CARBOHYD 973 973 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1007 1007 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1084 1084 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1103 1103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1114 1114 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1180 AA; 130808 MW; 8E5DA2BE02362EE4 CRC64;  
 Query Match 34.5%; Score 2149.5; DB 1; Length 1180;  
 Best Local Similarity 38.0%; Pred. No. 2e-131;  
 Matches 463; Conservative 242; Mismatches 405; Indels 107; Gaps 24;  
 QY 11 WALSLWPGTDTFNNMDTRKPRVIGSRATFAFGYTVQOQHDISGNKVLVVGAPLETNGYOKT 70  
 DB 17 WLLTILGFCYFNVVDVKNMSKSPQVEDMFGYTVQOYENBEGKVLGSLVGVGPQKART 76  
 QY 71 GGVYKCPVTHGN---CTKLNLG-RVTLNVSERKDNMLGLSLATNPKNDSFLACSPILWS 126  
 DB 77 GGVYKCPVGRERAMPCKVLDLPVNTSIPNVTIEKENMTFGSLVLTNP-NGGFLACGPPLA 135  
 QY 127 HECGSSVYTTGMCSSRVNSNFRSKVAPALORSCQYMDIVIVLDGSNSIYPMVEVQHFLI 186  
 DB 136 YRCGHLHYTTGICSDVSPFQVNSFAP-VQECSTQLDIVIVLDGSNSIYPMVESVIAFLN 194  
 QY 187 NILKFFYIGPQIQGVGVQYGEDVHVFHLDNYSRVKDVVEAASHIEQGGTETETAFQI 246  
 DB 195 DLLKRMDIGPKQTQVIGVGENVTHFNLNKYSSTEEVLVAANKIGRQGLQTMALGI 254  
 QY 247 BEARSEAF--QKGRKGAKKVMIVITDGHSDHSDPLEKVLQOSERNVTRYAVAVLYGN 304  
 DB 255 DTAKEAFTEARGARRGVKVMIVITDGHSDHNYRLKQVIOCCDENIQRFSLALGHYN 314  
 QY 305 RRGINPETPLNEIKYIASDPDDKHFFNVTDAAALXDIVDALGDRIFSLSGT-NKHETSPG 363  
 DB 315 RGNLSTEFVEIKSIASEPTKHFNFVSDALATVIVKALGERIFALEATADQAASPE 374  
 QY 364 LEMSGTGFSSHVVEDGVLLGAVDNGAVLKEFSAGKVIPLRSYLKEPPEELKNHCA 423  
 DB 375 MEMSGTGFSAHYSQDVMVLGAVDNGVVMQKQVAVNTPHNTTTFOTEPANMEPLAS 434  
 QY 424 YLGYTVTSVWSRQGRV-YVAGAPFNHTGKVLFTMNNRSLTHQAMRGQOIGSYFGS 482  
 DB 435 YLGYTVNS--ATIPGDVLYIAGPRNHTGVVIYKMDG-NINILQTLGGGQIGSYFGS 491  
 QY 483 EITSVDIDGQVTVLLVIGAPMYF-NEGREKGVVYVYELRQNRVYN-----GT 530  
 DB 492 VLTITDIDKDSYTDLLVIGAPMYGTEKEQGVVYVAVNTRFVQMSLEPIRQTCSS 551  
 QY 531 LKDSHSYO-----NARFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIPHGRGS 583  
 DB 552 LKONSTKENKNEPCGARFGTAIAAKDLNVDGVDVVGAPLEDHAGAVIYHGSCKT 611  
 QY 584 ILKTPKQITASELATGLQYFGCSIHGOLDINBGLIDLAVGALGNAILSRPVVQINA 643  
 DB 612 IREAVAQRIPSGGDGKTLKFFQSIHGEDMLNGDGLTDVTIGLGGALFWARDVAVVKV 671  
 QY 644 SLHPEPSKINIFHRDCKSRGATCAALPCTFPIFLAPHQFTTTVGIRYNATMDERYT 703  
 DB 672 TNHFEFNKNTQKNCVRVEGNETVCINATMCFHVKLSKEDSIYEADLOQYRVLDSLQI 731  
 QY 704 PRAHLEGGDRFTNRAVLLSSQELCERINHFVLDATDYVXPVTFSVESYSLDPOHGM 763  
 DB 732 SRSPSGTQERKIQRNITVRESE--CIRSHFVMLDKHDPQDSVRVTLDFNLTPDENG 789  
 QY 764 DQGWFTTLRVSVFVWNGCNEDEHCVPLAVLDARSDLPTAMEYQCVLRKPAQDCSAYLS 823  
 DB 790 DDLNPSVHEHIPPANOCGNKERCISDLTNAVST----- 824  
 QY 824 FDTTVFIISSTRQVRVAEATLENRGNAYSTVLNMQSANLQFASL--IQKEDSDGSIK 881  
 DB 825 -EKSLLIVKSQHDKNVSLTVKNKGDSAYNTRTVQVHSPNLI-FSGIEEIQKDCESN--- 880  
 QY 882 VNEERRLQKQVNCVNSYPPFRKAKVAVFLDSEFSKIFLHLELELAAGSDNERNDSKE 941

Db 881 -----QNTCKVGVFFRAGETVTKIFOFNTSHLSENAIHLATSDEEPLESJN 933  
 Qy 942 DNVAPLRFHLKYEADVLFTRSSLSHYEVLKASSLERY-----DGIGPPSPSCIFRIQNLGL 997  
 Db 934 DNEVNISIPVYEVGVGLQPYSSASEHHISVAANETIPEFINSINSTEEDIGNEINVFYTIKRGH 993  
 Qy 998 FPIHGMXXITPIATRSGRNLLKLDFTLDEANTSCNITWGNSTYRTPVTE--DLRRA 1055  
 Db 994 FPMPEQLSIFPNLTADGVPVLYPIGWSSSD--NVCN-----RPSLEDPPGINS 1043  
 Qy 1056 POLNEHNSDW-----SINCRLVENVQBEINPHLLGNLWLRSLKALKVKS 1100  
 Db 1044 KGMTSKSEVLKGTITQDCSSCGVATITCSLLPSDLQVNVSL--LKPFTIRAHFSS 1101  
 Qy 1101 KXIMMNAALQROFHPSPFIREBDSQIEFESKQB-DWQVPIWIVGSLTGLLLALL 1159  
 Db 1102 LNLTLRGLKSE-NSSLTLSSNRKRELAIQSKGLGRVPLWILLFAFAGLLMLL 1160  
 Qy 1160 VLALRKLGFFRGARRR 1176  
 Db 1161 ILALWKIGFFRKLKK 1177  
 [1]  
 RESULT 5  
 ITA2 MOUSE  
 ID ITA2 MOUSE STANDARD; PRT; 1178 AA.  
 AC Q62469; Q62163;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GPIa)  
 DE [Collagen receptor] (VLA-2 alpha chain) (CD49b).  
 GN ITGA2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;  
 RX MEDLINE=94363406; PubMed=8081889;  
 RA Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,  
 RA Damjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;  
 RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but  
 RT not virus binding."  
 RL Cell Adhes. Commun. 2:131-143(1994).  
 RN [2]  
 RP SEQUENCE OF 450-1178 FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=94355691; PubMed=7521231;  
 RA Wu J.E., Santoro S.A.;  
 RT "Complex patterns of expression suggest extensive roles for the alpha  
 RT 2 beta 1 integrin in murine development."  
 RL Dev. Dyn. 199:292-314(1994).  
 CC -1- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING  
 CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO  
 CC COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,  
 CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
 CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN  
 CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION  
 CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2  
 CC ASSOCIATES WITH BETA-1.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -1- SIMILARITY: Contains 1 VWFA domain.  
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z29987; CA82877.1; -;  
 DR EMBL; X75427; CA853178.1; -;  
 DR PIR; S44142; S44142.  
 DR HSSP; P17301; IAOX.  
 DR MGD; MGI:96600; Itga2.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWF A.  
 DR Pfam; PF01839; FG-GAP 3.  
 DR Pfam; PF00357; Integrin\_A; 1.  
 DR Pfam; PF00092; vwa; 1.  
 DR SMART; SM00191; Int\_alpha; 4.  
 DR SMART; SM00327; VWF; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS0234; VWFA; 1.  
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Platelet; Signal; Repeat; Calcium; Magnesium.  
 FT SIGNAL 1 26  
 FT CHAIN 27 1178  
 FT INTERIN ALPHA-2.  
 FT EXTRACELLULAR (POTENTIAL).  
 FT POTENTIAL.  
 FT CYTOPLASMIC (POTENTIAL).  
 FT FG-GAP 1.  
 FT FG-GAP 2.  
 FT VWFA.  
 FT FG-GAP 3.  
 FT FG-GAP 4.  
 FT FG-GAP 5.  
 FT FG-GAP 6.  
 FT FG-GAP 7.  
 FT POTENTIAL.  
 FT CA\_BIND 496 504  
 FT CA\_BIND 560 568  
 FT CA\_BIND 624 632  
 FT SITE 480 482  
 FT SITE 1154 1158  
 FT DISULFID 80 89  
 FT DISULFID 677 734  
 FT DISULFID 786 792  
 FT DISULFID 862 873  
 FT DISULFID 1016 1047  
 FT DISULFID 1052 1057  
 FT CARBOHYD 102 102  
 FT CARBOHYD 109 109  
 FT CARBOHYD 429 429  
 FT CARBOHYD 457 457  
 FT CARBOHYD 472 472  
 FT CARBOHYD 696 696  
 FT CARBOHYD 1054 1054  
 FT CARBOHYD 1071 1071  
 FT CARBOHYD 1078 1078  
 SQ SEQUENCE 1178 AA; 128925 MW; 1F194B9C0240F465 CRC64;  
 Query Match 30.7%; Score 1910.5; DB 1; Length 1178;  
 Best Local Similarity 36.2%; Pred. No. 6.6e-116;  
 Matches 442; Conservative 238; Mismatches 437; Indels 103; Gaps 31;

Qy 6 GLVVAWALSMPGFTD--TFNMDTKPRVIGSRRTAFPGYTVQOHDISGNKWLVVYGAPLE 63  
 Db 8 GALLIQLMLVQGIINCLAYNGLPKAIFGSPSEQFGYSVQQLTNGQWLLVGSFWS 67  
 Qy 64 TNGYQKTGDEVKCPV--THGNTKLN--GRVTLNVSVSRKKNRMLGLSLATNPKNDSFLA 120  
 Db 68 GFPENRMGDEVKCPVDLPATCEKLNQNSASISNWTBIKTMSLGLTLTRNPGTGGFLT 127  
 Qy 121 CSPLWSHECGSYVTTCMSRVNSFRSKTVAPALQRCOTYMDIVIVLDGNSNLYPWVE 180  
 Db 128 COPLMAHQCGNQYATGICSDVSPDFQLTSPPAVQACPSLVDVVVVCDEBSNLYPWEA 187  
 Qy 181 VQHFLINILKRFYIGGQIQGVQYGVQGVVHFEFLNDYRSVKDYVVEAASHIEQRGGT 240

Db 188 VKNPLVAVTGLDIPKPKTQVALLIQYANEPRIIFNLNDFETKEDMVQATSETRQHGDLT 247  
Qy 241 RTAFGLFARSEAFQK--GRKGAQVMIVITGESHDSPDLKLVICOSEEDNVTYAVA 298  
Db 248 NTRALFADYATSGRPGATKVMVVVITGESHDGSKLVITQCCNDLILFGIA 307  
Qy 299 VLGYINRGINPTEFLNEIKYIASDPDDKFFNVNTEAAALKTVDALGORIFSELTNKH 358  
Db 308 VLGYLNALDITKNEIKAIKIASTPTERYFFNVADEAALLKXAGTLGEOIFSEGTQV 367  
Qy 359 ETSFGLSEMSQTGSSHV--EDGVILGAVGAYDWNGAVLKETAGKVIPLRESYLKEPPE 416  
Db 368 GDNFQEMAGVGSADYAPQNDILMLGAVGAFDMSGLTVOETSHKPI-----FPR 418  
Qy 417 EL-----KNGHAYLCYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMHNNRSITIHQ 469  
Db 419 QAFDVLQDRHSSFLGYSV-AAISTEDGWHFVAGAPRANYTGOIVLYSVNKGQNVTVIQ 477  
Qy 470 AMGQQIGSGVGEISITVDIDGQVTVLVLGAPMFNE--GRERGKVVVELRONFVN 528  
Db 478 SHRGDQIGSGVGLSVLCSVDKDTITDVLVGAFTYNDLKKBEKGVLFTITKGLNQH 537  
Qy 529 GTLKDSHYQARFGSGIASVRLNODSYNDVVVGAPELDHAGAYIFHFGSGILKTP 588  
Db 538 QFLEGPCTGNARFGSAIALSDINDMDGNDVIVGSPVENSGAVIYNGHGTIRTKY 597  
Qy 589 KORITABELA--TGLVFGSGISGQDLNEDGLDLAVGALGNVILWSPVVOIKASLH 646  
Db 598 SOKILGNGAFRRHLPFGESLDGGLNGDSITDVSIGALGOVILWMSQSIADVAIEAL 657  
Qy 647 PPSKINIHRDCKRGRDATLAAFLCFTPIFLAPHFTTIVGIRVATMD---ERRY 702  
Db 658 FTEPKITILNKDAK-----ITLKLCPRAEP-RPAGQNNQVALLFNMTLDADGHSRV 708  
Qy 703 TPRAHLDGDDRFTNRAVLSSQGEICERINFHVLDTADYVKKPVTSEVSLDPDHGPM 762  
Db 709 TSGVFPRENERFLQKNWVNEVQK-CSEHHSITQKPSDVVNPLDLKRVDSLENPGTSPA 767  
Qy 763 LDDGWPTLAV-SVFPWANGNEBHCVPDLVLARSDLPTAMEVCQVLEKPAQDCSAYT 821  
Db 768 L-EAYSTVKTFSIFFYKESDGLCISDLILVQ-QLP-----804  
Qy 822 LSPDTTFFIESTRQVAVATEILNRCENAYSTVNLISQSANIOPASLIQKEDSGSIEC 881  
Db 805 -AIQTOSFIVSNOKRUTFSVLKNGESAYNTVLAERFSENLFASFSFMPVD---GTEV 860  
Qy 882 VNEERLQKV-CNVSYPPRAKAKVAFRLDSFSKSIPLHLEIELAGSDSNEDSTK 940  
Db 861 TCEVSSQKSVTCDVGPALKSEQQVTFINFPNLQNLQNLQNLQNLQNLQNLQNLQNLQ 918  
Qy 941 EDNVAPLRFHLKYEADVLFRSSSLSHYVKLN-----SSLERYDGGIPPFSCIFRIQ-NL 995  
Db 919 ADNSVSLTIPLYDABLHTRSTNINFEYISSDENAPSVIKSVEDIGPKF--IFSLKVT 976  
Qy 996 GLPFIHGMKMIPIATRSNRLKLDLDELTEA-NTSCNINWNSTEYTPV-----1048  
Db 977 GSAPVSNALVTIHPQVTEKNPLLYLTGTQDQAGDISC-----TABINPLKPTAPS 1031  
Qy 1049 ----EEDLRRAPOLNHSNDVSVINCNR-LVNPQEIHFLLGNLRLSLKALKYKSMKI 1103  
Db 1032 VSPKNEFRHTKELDCITSCSNITCWLKDLHKAIFYNVTWNRTFAASTPQT--V 1089  
Qy 1104 MVNAAORQFSPPIFREDDPSQIEFEISK-QEDMOVPWTIIVGSTLGLLLALLILVA 1162  
Db 1090 QLTAABEIDHTNPOLFVIEBNVAVTIPLMINKPEKAEVPTGVIGIIGILLAMTAG 1149  
Qy 1163 LRKLGPFRRARRRREPGLDP 1182  
Db 1150 LRKLGPFK--ROYKMGQNP 1167

RESULT 6

ITA2 BOVIN  
ID ITA2 BOVIN STANDARD; PRT; 1170 AA.  
AC P53710;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)  
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).  
GN ITGA2.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
SEQUENCE FROM N.A.  
RP MEDLINE=94193647; PubMed=7511592;  
RX Kamata T., Puzon W., Takada Y.;  
RT "Identification of putative ligand binding sites within I domain of  
integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";  
RL J. Biol. Chem. 269:9659-9663(1994).  
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,  
COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND B-CADHERIN. IT  
RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-E-R IN  
CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE  
EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
EXTRACELLULAR MATRIX.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2  
CC -!- ASSOCIATES WITH BETA-1.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L25886; AAB59255.1; -.  
DR FIR; L45914; L45914.  
DR HSSP; P17301; LAOX.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; VWF\_1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS0234; VWFA; 1.  
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.  
FT NON TER 1 1  
FT SIGNAL <1 18  
FT CHAIN 19 1170  
FT DOMAIN 19 1121  
FT DOMAIN 1122 1143  
FT TRANSMEM 1144 1170  
FT DOMAIN 1144 1170  
FT REPEAT 34 92  
FT REPEAT ? ?  
FT REPEAT 177 367  
FT DOMAIN ? ?  
FT REPEAT 423 475  
FT REPEAT 477 538  
FT REPEAT 540 599  
FT REPEAT 604 656  
FT REPEAT 604 656

|  |          |               |   |                                   |   |
|--|----------|---------------|---|-----------------------------------|---|
| FT   | CA_BIND  | 488           | 496   | POTENTIAL.                        |   |
| FT   | CA_BIND  | 552           | 560   | POTENTIAL.                        |   |
| FT   | CA_BIND  | 616           | 624   | POTENTIAL.                        |   |
| FT   | SITE     | 472           | 474   | CELL ATTACHMENT SITE (POTENTIAL). |   |
| FT   | SITE     | 1146          | 1150  | GFGR MOTIF.                       |   |
| FT   | DISULFID | 72            | 81  | BY SIMILARITY.                    |   |
| FT   | DISULFID | 669           | 726   | BY SIMILARITY.                    |   |
| FT   | DISULFID | 778           | 784   | BY SIMILARITY.                    |   |
| FT   | DISULFID | 854           | 865   | BY SIMILARITY.                    |   |
| FT   | DISULFID | 1008          | 1039  | BY SIMILARITY.                    |   |
| FT   | DISULFID | 1044          | 1049  | BY SIMILARITY.                    |   |
| FT   | CARBOHYD | 94            | 94  | N-LINKED (GLCNAC. . .)            | (POTENTIAL).                            |
| FT   | CARBOHYD | 101           | 101   | N-LINKED (GLCNAC. . .)            | (POTENTIAL).                            |
| FT   | CARBOHYD | 332           | 332   | N-LINKED (GLCNAC. . .)            | (POTENTIAL).                            |
| FT   | CARBOHYD | 421           | 421   | N-LINKED (GLCNAC. . .)            | (POTENTIAL).                            |
| FT   | CARBOHYD | 449           | 449   | N-LINKED (GLCNAC. . .)            | (POTENTIAL).                            |
| FT   | CARBOHYD | 464           | 464   | N-LINKED (GLCNAC. . .)            | (POTENTIAL).                            |
| FT   | CARBOHYD | 688           | 688   | N-LINKED (GLCNAC. . .)            | (POTENTIAL).                            |
| FT   | CARBOHYD | 748           | 748   | N-LINKED (GLCNAC. . .)            | (POTENTIAL).                            |
| FT   | CARBOHYD | 945           | 945   | N-LINKED (GLCNAC. . .)            | (POTENTIAL).                            |
| FT   | CARBOHYD | 1063          | 1063  | N-LINKED (GLCNAC. . .)            | (POTENTIAL).                            |
| FT   | CARBOHYD | 1070          | 1070  | N-LINKED (GLCNAC. . .)            | (POTENTIAL).                            |
| FT   | VARIANT  | 580           | 580   | G -> V.                           |   |
| FT   | VARIANT  | 588           | 588   | R -> K.                           |   |
| FT   | VARIANT  | 725           | 725   | R -> S.                           |   |
| SQ   | SEQUENCE | 1170 AA;      | 128929 MW;  | SECEFLCSF2448FB1                  | CRC64;                                  |
| Query Match  |          |               |   |                                   | 30.6%; Score 1901.5; DB 1; Length 1170; |
| Best Local Similarity  |          |               |   |                                   | 36.3%; Pred. No. 2.5e-115;              |
| Matches 437; Conservative 238; Mismatches 437; Indels 91; Gaps 30; |          |               |   |                                   |   |
| QY   | 23       | FMOTRKERVIPG  | STAFYVTVQOHDISGNKLVVWGLAPLETNGYQKTDVYKCPV--IH     | 80                                |   |
| Db   | 19       | YVNGVLPKAFIPG | SPSEOGFVAVQOFINPKGNWLVGSPWGFPPKRMGDVYKCPVDLST     | 78                                |   |
| QY   | 81       | GNCTKLNIGRVT  | -LSNVSERKDNMELGLSLATPNKDNSFLACSPLWSHECCSSYYTTCMC  | 139                               |   |
| Db   | 79       | TTCEKLNLOTSM  | NVTENKTNMSLGLTLTRNVTGCGFLCGPLMAQCCSQYYTTCVC       | 138                               |   |
| QY   | 140      | SRVNSNFRPSKT  | VPALQRCQYMDIVIVLDSNSIYPWVEVQHPLINILKXFIYGPQI      | 199                               |   |
| Db   | 139      | SDVSPDFQLT    | RFAPAVQTCPSFIDVWVVCDESNIYPWDVAVXNFKLFVQGLDGPTKT   | 198                               |   |
| QY   | 200      | CVGVVOYGEDV   | VHEPHLNDYSVKDVVEAASHIEORGCTETATFGIEFARSEAPQ--KG   | 257                               |   |
| Db   | 199      | QNGLIQYANPR   | VFNLTNFKSDEMIXATSTQTFQYGGDLTNTFKAIQTARDTAYSTAG    | 258                               |   |
| QY   | 258      | GRKGAKVMIVIT  | DGESHDSPLEKVIQOSRDNVTRYAVAVLYGYNRRGINPETFLNEI     | 317                               |   |
| Db   | 259      | GRPGATKVMVV   | TDGESHGSKLKAVIDQCNKDNILRFGIAVGLYLNALOTKULIKEI     | 318                               |   |
| QY   | 318      | KYIASDPDDKH   | FFNTDEAALKDVIDALGDRIPSLKCTNKTSTGLKMSQTSGRSSHW     | 377                               |   |
| Db   | 319      | KALASIPTRHF   | FNVSDEADLEKAGTIGRQIFSIIEGTVQGDNPQMEMSQVGSFSAEYSP  | 378                               |   |
| QY   | 378      | DG--VLLGAVG   | AYDNNGAVLKETSAGKVIPLRESYLKEPFBEL---KNHGAYLYVTTSV  | 432                               |   |
| Db   | 379      | QNNILMUGAVG   | AYDWSGTIVQKTPHGLI-----PSKQAFQILQDRNHSSYLYGSVAS-   | 432                               |   |
| QY   | 433      | VSSROGRVYV    | GAPFNHTGKVLITMNNRSLTHQAMRQOIGSYFGSITSVDIDGD       | 492                               |   |
| Db   | 433      | 1STGNSVHFV    | GAPANTYGTQIVLYSVNENGNTVIOQRQDQIGSYFGSLCAVDVNDK    | 492                               |   |
| QY   | 493      | GYTDVLLGAP    | YFNE-CRERKGVVYVELRNRFPVYNGTLKDSHYQARFGSSIASVRD    | 551                               |   |
| Db   | 493      | TITDVLGAP     | YMYMDLKEEGRVYLFITIKILNWHQFLGPNGLNARFGSAIALSD      | 552                               |   |
| QY   | 552      | LNQDSYNDV     | VVGAPLEDNHAGAIYIFHGRFSILKTPKQITASELA--TGLQVFGCSIH | 609                               |   |
| Db   | 553      | INMDGFNDV     | IGSPLENQSGAVIYNGHEGMIRLRYSQKILGSDRAFSSHLQVFGRLD   | 612                               |   |
| QY   | 610      | GOLDLNEGL     | IDLAVGALGNVILWSRPVVOINASLHFPBPSKINIHFHDCRSGSDATCL | 669                               |   |

|          |   |  |  |                              |     |
|----------|---|--|--|------------------------------|-----|
| Db       | 613   | GYGLDNGSD                                  | ITDVSVGAFQVQVLSQS                                    | IADVSVDASFTPKKILLNKNAE-----I | 664 |
| QY       | 670   | AALCFPTPI                                  | FLAPHFQTTTGIYRNATMDERRYT----                         | PRAHLEGGDRFTNRAVLSSG         | 725 |
| Db       | 665   | KLKLCFSAK                                  | P-RPTNQNNQVAIVNITIDEDQSSSRVISRGLFKENNERCLQKTMIVSQ    | 723                          |     |
| QY       | 726   | QELCERINF                                  | VLDTADYVVKPVPFSVYSLEDDPHGPMLDGWPPTTLRV-SYFFWNGCND    | 784                          |     |
| Db       | 724   | QR-CSEYII                                  | IIHQEPSDIIISPLNLCMNISLENPGTNPAL-EAYSETVAVPSIPFHKDCDD | 781                          |     |
| QY       | 785   | EHCVPDVL                                   | DARSDLPTAMEYCORVLKPKAQDCSAYTSLSDTTVFIIESTRQRAVENTL   | 844                          |     |
| Db       | 782   | GVCIISDL                                   | VLNVQ-QLPATQO-----QP-----                            | FIVSNQMKRLTFSVOL             | 818 |
| QY       | 845   | ENRGENAY                                   | STVLNISQSANLOFASLIQCKSDGSECVNBERRLOKV-CNVSYPPFFRAK   | 903                          |     |
| Db       | 819   | KNKESAYN                                   | TEIVVDFSENLFASNSMPVD---GTEVTCQIASSQKSVTCNVGYPAUKSK   | 875                          |     |
| QY       | 904   | AKVAFRLD                                   | SFBSKSIIFLHLIELELAAGSDNERDSTEDNVAPLRFHLKTEADVLPTRSS  | 963                          |     |
| Db       | 876   | QCVTFIN                                    | FDENLQNLQNASISPRALSESQENMA--DMSVNLKLSLLYDAEIHTRST    | 933                          |     |
| QY       | 964   | SLSHYEVL                                   | ---NSSLERVDGIGPPFCIFRIQ-NLGLPPIHGMWMMKITIPIATRSNR    | 1018                         |     |
| Db       | 934   | NINFEV                                     | SLDGNSSVSHSFEDIGPKF--LFSIKVTGTSVPSVSNASVIIHIPOYTKQNP | 991                          |     |
| QY       | 1019  | LLKLRLD                                    | FLTDQA-NTSCNIMWNSTEYRPTV-----BEDLRAPQLNHNNSDVS       | 1067                         |     |
| Db       | 992   | LWYLTGVH                                   | TQAGDISC-----EAEINPLXIGOTSSSVFSKSNFRHIKELNCRFASCSN   | 1046                         |     |
| QY       | 1068  | INCNIR                                     | -LVPNQENFHLGNLRLSKALKYKSKIMVNAALQORQHSPPIFEREPSR     | 1126                         |     |
| Db       | 1047  | IMCWRLQ                                    | VQKGYFLNVSTRIMNGTFAASTFQF--VOLTRAABIDTNPQIYVIENTV    | 1104                         |     |
| QY       | 1127  | QIEFEISK                                   | -QEDMQVPIIWIIVGSTLGGLLALLALVALRLKLGFFRSARRRBPGLDPTPK | 1185                         |     |
| Db       | 1105  | TIPLTIMK                                   | PEKVEVPTGVIGSVIAGILLALLVALLWKLGPPPKRYEKMAKNPDETDE    | 1164                         |     |
| QY       | 1186  | VLE  | 1188   |                              |     |
| Db       | 1165  | TTE  | 1167   |                              |     |
| RESULT 7 |   |  |  |                              |     |
| IT       | IT22_HUMAN  | STANDARD;                                  | PRT;   | 1181 AA.                     |     |
| AC       | PI7301;   |  |  |                              |     |
| DT       | 01-AUG-1990   | (Rel. 15, Created)                         |  |                              |     |
| DT       | 01-AUG-1990   | (Rel. 15, Last sequence update)            |  |                              |     |
| DT       | 10-OCT-2003   | (Rel. 42, Last annotation update)          |  |                              |     |
| DE       | Integrin alpha-2 precursor  | (Platelet membrane glycoprotein Ia) (GPIa) |  |                              |     |
| DE       | (Collagen receptor)   | (VLA-2 alpha chain) (CD49b).               |  |                              |     |
| GN       | ITG22.  |  |  |                              |     |
| OS       | Homo sapiens (Human)  |  |  |                              |     |
| OC       | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;     |  |  |                              |     |
| OC       | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.            |  |  |                              |     |
| OX       | NCBI_TaxID=9606;  |  |  |                              |     |
| RN       | [1]   |  |  |                              |     |
| RP       | SEQUENCE FROM N.A. AND SEQUENCE OF 30-44.                             |  |  |                              |     |
| RC       | TISSUE=Endothelial cells;   |  |  |                              |     |
| RX       | MEDLINE=89308879; PubMed=2545729;                                     |  |  |                              |     |
| RT       | Takada Y., Hemler M.E.  |  |  |                              |     |
| RT       | "The primary structure of the VLA-2/collagen receptor alpha 2 subunit |  |  |                              |     |
| RT       | (platelet GPIa): homology to other integrins and the presence of a    |  |  |                              |     |
| RT       | possible collagen-binding domain."                                    |  |  |                              |     |
| RL       | J. Cell Biol. 109:397-407 (1989).                                     |  |  |                              |     |
| RN       | [2]   |  |  |                              |     |
| RP       | SEQUENCE FROM N.A.  |  |  |                              |     |
| RA       | Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,     |  |  |                              |     |
| RA       | Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;                        |  |  |                              |     |
| RL       | Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.               |  |  |                              |     |
| RN       | [3]   |  |  |                              |     |
| RP       | X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.                     |  |  |                              |     |



RX MEDLINE=98019223; PubMed=9353312;  
 RA Emsley J., King S.L., Bergelson J.M., Liddington R.C.;  
 RT "Crystal structure of the I domain from integrin alpha2beta1.";  
 RL J. Biol. Chem. 272:28512-28517(1997).  
 RN (4)  
 RP VARIANT HPA-5 (BR).  
 RX MEDLINE=94043762; PubMed=7901236;  
 RA Santoso S., Kaib R., Walka M., Kiefel V., Mueller-Eckhardt C.,  
 RA Newman P.J.;  
 RT "The human platelet allantoicins Br(a) and Br(b) are associated with a  
 RT single amino acid polymorphism on glycoprotein Ia (integrin subunit  
 RT alpha 2).";  
 RL J. Clin. Invest. 92:2427-2432(1993).  
 RN [5]  
 RP VARIANT GLU-534.  
 RX MEDLINE=20206009; PubMed=10744142;  
 RA Kroll H., Gardemann A., Fechter A., Haberbosch W., Santoso S.;  
 RT "The impact of the glycoprotein Ia collagen receptor subunit A1648G  
 RT gene polymorphism on coronary artery disease and acute myocardial  
 RT infarction.";  
 RL Thromb. Haemost. 83:392-396(2000).  
 CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,  
 CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT  
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-G-E-R IN  
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER  
 CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE  
 CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
 CC EXTRACELLULAR MATRIX.  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2  
 CC associates with beta-1. Interacts with HPSS.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -!- POLYMORPHISM: Position 534 is associated with platelet-specific  
 CC allantoicins HPA-5 (BR). HPA-5A/BR(A) has Lys-534 and HPA-5B/BR(B)  
 CC has Glu-534. HPA-5B is involved in neonatal alloimmune  
 CC thrombocytopenia (NAIT or NATP). The K534E polymorphism may play a  
 CC role in coronary artery disease (CAD).  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD49b entry;  
 CC WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd49b.htm".  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; X17033; CAA34894.1; -.  
 CC EMBL; AF512556; AAM34795.1; -.  
 CC PIR; A33998; A33998.  
 CC PDB; IAOX; 25-NOV-98.  
 CC PDB; IDZI; 02-AUG-01.  
 CC Genew; HGNC:6137; ITGA2.  
 CC MIM; 192974; -.  
 CC GO; GO:0008305; C:integrin complex; TAS.  
 CC GO; GO:0005886; C:plasma membrane; TAS.  
 CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.  
 CC GO; GO:0005518; F:collagen binding; TAS.  
 CC GO; GO:0007596; F:blood coagulation; TAS.  
 CC GO; GO:0007160; P:cell-matrix adhesion; TAS.  
 CC GO; GO:0007397; P:histogenesis and organogenesis; TAS.  
 CC InterPro; IPR000413; Integrin\_alpha.  
 CC InterPro; IPR002035; VWFA.  
 CC Pfam; PF01839; FG-GAP; 3.  
 CC Pfam; PF00357; Integrin\_A; 1.  
 CC Pfam; PF00092; vwa; 1.  
 CC SMART; SM00191; int\_alpha; 5.  
 CC SMART; SM00327; VWFA; 1.

DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS50234; VWFA; 1.  
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;  
 KW 3D-structure.  
 FT SIGNAL 1 29  
 FT CHAIN 30 1181 INTEGRIN ALPHA-2.  
 FT DOMAIN 30 1132 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1133 1154 POTENTIAL.  
 FT DOMAIN 1155 1181 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 1155 1161 INTERACTION WITH HPSS.  
 FT REPEAT 45 103 FG-GAP 1.  
 FT REPEAT ? ? FG-GAP 2.  
 FT DOMAIN 188 378 VWFA.  
 FT REPEAT 378 433 FG-GAP 3.  
 FT REPEAT 434 486 FG-GAP 4.  
 FT REPEAT 488 549 FG-GAP 5.  
 FT REPEAT 551 610 FG-GAP 6.  
 FT REPEAT 615 667 FG-GAP 7.  
 FT CA\_BIND 499 507 POTENTIAL.  
 FT CA\_BIND 563 571 POTENTIAL.  
 FT CA\_BIND 627 635 POTENTIAL.  
 FT SITE 1157 1161 GFPR MOTIF.  
 FT DISULFID 83 92 BY SIMILARITY.  
 FT DISULFID 690 737 BY SIMILARITY.  
 FT DISULFID 789 795 BY SIMILARITY.  
 FT DISULFID 865 876 BY SIMILARITY.  
 FT DISULFID 1039 1050 BY SIMILARITY.  
 FT DISULFID 1055 1060 BY SIMILARITY.  
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 343 343 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 440 440 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 455 455 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 699 699 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1057 1057 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1081 1081 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARIANT 534 534 K -> E (IN ALLOANTIGEN HPA-5B;  
 FT dbSNP:1801105).  
 FT /FTID=VAR\_003977.  
 FT TURN 170 171  
 FT STRAND 173 180  
 FT TURN 183 184  
 FT HELIX 188 199  
 FT TURN 200 201  
 FT STRAND 204 204  
 FT STRAND 206 207  
 FT STRAND 209 216  
 FT STRAND 220 224  
 FT TURN 226 228  
 FT HELIX 232 240  
 FT TURN 241 241  
 FT HELIX 252 262  
 FT TURN 263 284  
 FT HELIX 266 288  
 FT TURN 269 269  
 FT STRAND 275 282  
 FT HELIX 289 291  
 FT HELIX 292 301  
 FT TURN 302 303  
 FT STRAND 304 311  
 FT HELIX 313 317  
 FT TURN 318 319  
 FT HELIX 323 330  
 FT TURN 331 332  
 FT HELIX 337 340  
 FT STRAND 341 344  
 FT HELIX 347 353  
 FT HELIX 354 362  
 FT TURN 363 363  
 SQ SEQUENCE 1181 AA; 129295 MW; 7E1B7ED968A94070 CRC64;

Query Match 29.9%; Score 1863; DB 1; Length 1181;  
Best Local Similarity 35.6%; Pred. No. 8.1e-113;  
Matches 432; Conservative 241; Mismatches 433; Indels 106

[illegible]

|    |      |   |       |     |
|----|------|---|-------|-----|
| Qy | 994  | NIGLFIPIHGMKMTIPIATRSNRLIKLRDLFTDEA-NTSCNIWGNSTYRPTP        | ----- | 104 |
| Db | 978  | TTGSGVPSMATVLIHPIQYTKKCNPLMYLTGVGTDRAGDISCNADINPLKIGTSGSVSF | 1037  |     |
| Qy | 1048 | VEEDLRAPOLNHSNDVVSINCINLVPNQEIFHLG-----NLWLRSLKALKYK        | 1099  |     |
| Db | 1038 | KSENFHRTKELNCRTASCNSVTCLK-----DVEMKGEYFVNVTIRWNGTFASSTPQ    | 1090  |     |
| Qy | 1100 | SMKIMNVALQROFHSGPFIREDPSQRQIEFEISK-OBQWQVPIWIVGSTIGGLLLAL   | 1158  |     |
| Db | 1091 | T--VQLTAAARINTYNPETVIEDNTVTIPLMIKPKDEKAEVPTGVITIGSIAGILLLLA | 1148  |     |
| Qy | 1159 | LVLALRKLGPFRR   | 1170  |     |
| Db | 1149 | LVALWKLGPFK   | 1160  |     |

RESULT 8

ITAL\_HUMAN

ID

ITAL\_HUMAN

STANDARD;

PRT;

1170 AA.

AC

P30701; 043746;

DT

01-FEB-1991 (Rel. 17, Created)

DT

01-NOV-1995 (Rel. 32, Last sequence update)

DT

10-OCT-2003 (Rel. 42, Last annotation update)

DE

Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1 alpha chain) (Leukocyte function associated molecule 1, alpha chain)

DE

(CD11a)

GN

ITGAL OR CD11A.

OS

Homo sapiens (Human).

OC

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX

NCBI\_TaxID=9606;

RN

[1]

RP

SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.

RA

MEDLINE=89139587; PubMed=2537322;

RA

Larson R.S., Corbi A.B., Berman L., Springer T.;

RT

"Primary structure of the leukocyte function-associated molecule-1 alpha subunit: an integrin with an embedded domain defining a protein superfamily.";

RL

J. Cell Biol. 108:703-712(1989).

RN

[2]

RP

SEQUENCE FROM N.A. (ISOFORM 2).

EX

MEDLINE=99425270; PubMed=10493829;

RA

Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R., Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;

RT

"Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 1cp and 16q.";

RL

Genomics 60:295-308(1999).

RN

[3]

RP

X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.

EX

MEDLINE=96036057; PubMed=7479767;

RA

Qu A., Leahy D.J.;

RT

"Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L beta 2) integrin.";

RL

Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281(1995).

RN

[4]

RP

X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.

EX

MEDLINE=96398682; PubMed=8805579;

RA

Qu A., Leahy D.J.;

RT

"The role of the divalent cation in the structure of the I domain from the CD11a/CD18 integrin.";

RL

Structure 4:931-942(1996).

RN

[5]

RP

X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.

EX

MEDLINE=99425289; PubMed=10493852;

RA

Kallen J., Weizenbach K., Ramagge P., Geyl D., Kriwacki R., Legge G., Cottens S., Weitz-Schmidt G., Hommel U.;

RT

"Structural basis for LFA-1 inhibition upon luvastatin binding to the CD11a I-domain.";

RL

J. Mol. Biol. 292:1-9(1999).

CC

!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,



CC ICAM3 AND ICAM4. IT IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA  
 CC INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL  
 CC MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES  
 CC AND MONOCYTES.  
 CC  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L  
 CC ASSOCIATES WITH BETA-2.  
 CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Bvert-Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=P20701-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=P20701-2; Sequence=VSP\_002738;  
 CC Note=No experimental confirmation available;  
 CC TISSUE SPECIFICITY: LEUKOCYTES.  
 CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -1- SIMILARITY: Contains 1 VWFA domain.  
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD11a entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11a.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Y00796; CAA68747.1; --  
 CC EMBL; AC002310; AAC31672.1; --  
 CC FIR; S03308; S03308.  
 CC PDB; 1LFA; 29-JAN-96.  
 CC PDB; 1ZON; 07-DEC-96.  
 CC PDB; 1ZOO; 07-DEC-96.  
 CC PDB; 1ZOS; 07-DEC-96.  
 CC PDB; 1CQF; 07-AUG-00.  
 CC PDB; 1DGO; 03-FEB-00.  
 CC PDB; 1MCN; 28-JAN-03.  
 CC PDB; 1MQ9; 14-JAN-03.  
 CC PDB; 1MQ9; 14-JAN-03.  
 CC PDB; 1MQA; 14-JAN-03.  
 CC Genem; HGNC:6148; ITGAL.  
 CC MIM; 153370; --  
 CC GO; GO:0008305; C:integrin complex; TAS.  
 CC GO; GO:0006928; P:cell motility; TAS.  
 CC InterPro; IPR000413; Integrin\_alpha.  
 CC InterPro; IPR002035; VWF\_A.  
 CC Pfam; PF01839; FG-GAP\_3.  
 CC Pfam; PF00357; integrin\_A; 1.  
 CC Pfam; PF00092; vwa; 1.  
 CC PRINTS; PR01185; INTEGRINA.  
 CC PRINTS; PR00453; VWFADOMAIN.  
 CC SMART; SM00191; Int\_alpha; 4.  
 CC SMART; SM00327; VWA; 1.  
 CC PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 CC PROSITE; PS00234; VWFA; 1.  
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 CC Signal; 3D-structure; Magnesium; Calcium; Repeat;  
 CC Alternative splicing.  
 CC SIGNAL 1 25  
 CC CHAIN 26 1170 INTEGRIN ALPHA-L.  
 CC DOMAIN 26 1088 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 1089 1112 POTENTIAL.  
 CC DOMAIN 1113 1170 CYTOPLASMIC (POTENTIAL).  
 CC REPEAT 42 91 FG-GAP 1.  
 CC REPEAT 92 149 VWFA.  
 CC DOMAIN 170 349 FG-GAP 2.  
 CC REPEAT ? ? FG-GAP 3.  
 CC REPEAT 401 455 FG-GAP 4.  
 CC REPEAT 457 516 FG-GAP 5.

FT REPEAT 518  
 FT REPEAT 578  
 FT CA\_BIND 468  
 FT CA\_BIND 530  
 FT CA\_BIND 598  
 FT SITE 1115  
 FT SITE 1119  
 FT DISULFID 73  
 FT DISULFID 80  
 FT DISULFID 111  
 FT DISULFID 129  
 FT DISULFID 653  
 FT DISULFID 707  
 FT DISULFID 771  
 FT DISULFID 845  
 FT DISULFID 998  
 FT DISULFID 1021  
 FT DISULFID 1052  
 FT CARBOHYD 65  
 FT CARBOHYD 89  
 FT CARBOHYD 188  
 FT CARBOHYD 649  
 FT CARBOHYD 670  
 FT CARBOHYD 726  
 FT CARBOHYD 730  
 FT CARBOHYD 862  
 FT CARBOHYD 885  
 FT CARBOHYD 897  
 FT CARBOHYD 1060  
 FT CARBOHYD 1071  
 FT CARBOHYD 954  
 FT VARSPLIC  
 FT  
 FT CONFLICT 214  
 FT CONFLICT 660  
 FT STRAND 155  
 FT STRAND 162  
 FT TURN 164  
 FT TURN 166  
 FT HELIX 169  
 FT TURN 185  
 FT TURN 188  
 FT STRAND 191  
 FT STRAND 198  
 FT STRAND 202  
 FT STRAND 208  
 FT HELIX 217  
 FT TURN 221  
 FT TURN 222  
 FT STRAND 229  
 FT STRAND 233  
 FT TURN 244  
 FT TURN 247  
 FT HELIX 250  
 FT TURN 252  
 FT TURN 254  
 FT STRAND 256  
 FT HELIX 274  
 FT TURN 277  
 FT STRAND 280  
 FT STRAND 286  
 FT HELIX 288  
 FT HELIX 293  
 FT TURN 298  
 FT TURN 300  
 FT HELIX 307  
 FT STRAND 311  
 FT TURN 318  
 FT TURN 319  
 FT TURN 328  
 FT TURN 329  
 FT TURN 330  
 SQ SEQUENCE 1170 AA; 128819 MW; 39A7AP92EF286FC0 CRC64;

Query Match 18.4%; Score 1146.5; DB 1; Length 1170;  
 Best Local Similarity 29.6%; Pred. No. 2.6e-66;  
 Matches 373; Conservative 200; Mismatches 460; Indels 229; Gaps 56;  
 QY 8 VVAVALLSLWPGF-----TDTFNMTRKPRVPGSRTA-PFGYTVQOHDISGNKWLVGAP 61  
 Db 6 ITVMALLSGFFFPASSVNLVGRGARSPPGAGRHFGYRVLQ----VGN-GVIVGAP 61  
 QY 62 LFTNGYQKTDVYKCPVTHGNTKLNGLRVTL--SNVSEKDNKRELGLSLATNPKNDSFL 119  
 Db 62 GEGN---STGSLYQQQSGTGHCLP-----VTLRGSNYTSK-----YLGHTLATDPTDGSIL 109

QY 120 ACSPLWSECGSYTTGMSRVNSNFRFSKTVPA-PALQRC-QTYMDIVIVLDGSGSYTYP 177  
DB 110 ACDGSLSTCDQNTYLSGLCYLFRQNLQGMQGRPGQECIKGNVDVFLFDGMSLQSP 169  
QY 178 --WVEVQHPLNLIKFYIGPOQIQGVVQYGEDVYHBFHNDYSVDVVEAASHIEQR 235  
DB 170 DEFQKILDFMDKWK--LSNTSYQFAAVQFSTYKTEFDYVYKRPDALLKXVHM 227  
QY 236 GGTETRTAFGIEFARSEAFQK--GGRGKAKVMIVITGSEHSDSDPLEKVTQOQSRDNTV 293  
DB 228 -LLLTNTGAINVATEVREBELGARPDATKVLIIITDGEAVDSGNIDAA-----KDLI 280  
QY 294 RYAVAVLGYNNRRGINPETFLNEIKYIASDPDKRHFNNVTDE-AALKDVIDALGRIFSL 352  
DB 281 RYIIGI-GSHFTKSSQET---LHKFASKPASE-FVKILDTFEKLDFTLOKXIVVI 334  
QY 353 EGTNENE-TSFGLEMSQTCFSSHVEDGVLLGAVGAYDNGAVLKETSAGKVIPIR---- 407  
DB 335 EGTSKQDLSFNMELSSGSIADLSRGAHVAVGAVKOW-----AGGLDLKADLQ 385  
QY 408 -ESYLKEPPEELKNEHGAVLGYTVTVSVSROGRVYVAGAPRNNHTGKVLFTM-----HN 461  
DB 386 DTFPIGNEPLTEVRAGYLGTVTWLPSPQKTSILLASGAPRYOHMGRVILLFOEPGGGHV 445  
QY 462 NRSLLTHOARCOQICSGYFSGSITSVDIDGQVTDVLLVGAPMYNEGRERKGVYVIELR 521  
DB 446 SOVQTIH-----GTOIGSYFGELCGVDVDQGETELIGAPLFYGEOR-GGRVFIYQRR 500  
QY 522 QNRFYVNGTLKDSHYSQNAFCSSIASVRDLNQDSVNDVVCAPLEDNHAGAIYIHFHGR 581  
DB 501 QLGFEVSELOQDPGYPIGRFGEAITALTDINGDLGVAVAGAPLEEQ--GAVYIFNGRH 558  
QY 582 GSILTKPQRITASELATNGLQVFGCSIHGQDLNEDGLIDLAVAGNALVILWSPPVQVI 641  
DB 559 GGLSPQSORIETGTVLSGIQWFGERSIHGKDLGGLADVAVAGSOMVILSSPPVDM 618  
QY 642 NASLHPEPSKINPHRDK-----RSGRDATCLAAFLCFTPIELAPHEOTTVV-GIRY 693  
DB 619 VTLMGFSFPAIIPVHEVECSYSTSNMKKEGVNIT-----ICFOIKSLYPOFGRLVANITY 673  
QY 694 NATMDERRYTPRAHLDEGGDRFTNRAVLLSSQBELCERINFIV-LDTADYKVPVFSVEY 752  
DB 674 TLQLDGHR--TRRGLFPGRGHELRNIAVTGMS--CTDPSFHPVQVQDLISPINVSLNF 731  
QY 753 SLEDDPHGP-----MLDDGHPPTILRV-----VFFWNGCNEDEHCVDPDLVLDARSDLPT 801  
DB 732 SLWBEEGTPDQRAQCKDIPPIRLPSLSHSETWEIPFKNCBGDKCEANL----- 781  
QY 802 AMEYQVRLKPAQDCSAVTLTDFDTVTFTIESTRQVAVEATLENKGENAYSTVLNISQS 861  
DB 782 -----RVSPSPARSALRLTAFS-----LSVELSLNLEDAVWVQDLHFFP 824  
QY 862 ANLQFASL-IQKEDSDGSIET--VNERRLQKV--CNVSYFFFRKAKVAPRL----- 910  
DB 825 PGLSFRKEMLPKHSQIPVSCBELPEESRLSRLSCNVSSPIFKAGHSVALQMMFNFLV 884  
QY 911 DSEPSKSIPLHLEIELAAGSDSNEEDSTKEDONAPLAFHLKYADVL----- 958  
DB 885 NSSWGDSEVH-----ANVTCNEDSDLLENSATTIPIPIPNILIQOEDSTLVS 938  
QY 959 FTRSSLSH-----YEVKLNS-----LERYDGIQPPFCIFRIQNLGLFFPIHGM 1005  
DB 939 FTKPKGIHQVGMVQVRIQPSIHDNHIPTLEAVGVQVQ-----PSEG--- 982  
QY 1006 KITIPATSGNRLKLRDLTDEANTSCNIGNSTEYRPTVE-EDLRRAPQNLHNSD 1064  
DB 983 ----PITH-----MSVQME-PPVPCVHYEDLERLPDAAABCPIC 1015  
QY 1065 VVSINCIRLVNQBINPHILGNML-----RSLKALKYKSMKIMVNAALQRPQSPFI 1118  
DB 1016 GAFRCFV--VFRQILVQVITLVLNGEIEASSMFL-CSLSISFNS--KHFF---L 1067  
QY 1119 FREEDPSQIEPSEIKQEDMQVPIWIIIVGSTTGGLLLLALLVLRKLGFF-RSARRRE 1177

DB 1068 YGSNASLAQVMKVVDVYVYBKQM-LYLVLSIGIGGULLLLLFIVLYKVGFFKRLKXOME 1126  
QY 1178 PG 1179  
DB 1127 AG 1128

RESULT 9

ID ITAD\_HUMAN STANDARD; PRT; 1162 AA.

AC Q13349; Q15575; Q15576;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).  
GN ITGAD.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RX MEDLINE=96111956; PubMed=8777714;  
RA Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,  
RA Staunton D.E., Gallatin W.M.;  
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-3".  
EL J. Biol. Chem. 275:9959-9969(2000).  
RN [2]  
RP SEQUENCE OF 1-235 FROM N.A.  
RX MEDLINE=20187620; PubMed=10722744;  
RA Noti J.D., Johnson A.K., Dillon J.D.;  
RT "Structural and functional characterization of the leukocyte integrin gene CD11d. Essential role of Spi and Sp3".  
EL J. Biol. Chem. 275:9959-9969(2000).  
RN [3]  
RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.  
RX MEDLINE=96257236; PubMed=8666289;  
RA Wong D.A., Davis E.M., LeBeau M., Springer T.A.;  
RT "Cloning and chromosomal localization of a novel gene-encoding a human beta 2-integrin alpha subunit".  
EL Gene 171:291-294(1996).  
RN [4]  
RP INTERACTION WITH VCAM1.  
RX MEDLINE=99059842; PubMed=9841932;  
RA Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,  
RA Hoffman P.A., Staunton D.E., Bochner B.S.;  
RT "alpha beta 2 integrin is expressed on human eosinophils and functions as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-1)".  
EL J. Exp. Med. 188:2187-2191(1998).  
RN [5]  
RP INTERACTION WITH VCAM1.  
RX MEDLINE=99370002; PubMed=10438935;  
RA Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,  
RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;  
RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a binding interface between I domain and VCAM-1".  
EL J. Immunol. 163:1984-1990(1999).  
CC -!- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-BORNE PATHOGENS, PARTICULATE MATTER, AND SENSIBLE ERYTHROCYTES FROM THE BLOOD.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D ASSOCIATES WITH BETA-2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.



QY 883 --NEERLQKQVNSYPPFRKAKVAFRLDSEFS--KSLFLHLELELAAGSDSNERDST 939  
 Db 849 VPTEDBGLASSRCVNHPIFHEGSGNCTFVTDVSVKATLGRMLMRASENKASS 908  
 QY 940 KEDNVAPLRFHLKYADVLFTRSSLSHY-----EVKNSLSRYDGIQPPFSCIPRI 992  
 Db 909 KA--TFQELPKVYAVYTMISRQSESTKFNATSDKKOKAEHRY-----RV 955  
 QY 993 QNLGLPPIHGMKTIPIATRSNRLKLRDLPLDDEANTS--CNWNGNSTEYRPTPVEE 1050  
 Db 956 NNLQSDRL-AISINFWFVLLNG-----VAVDVWVEAPSOQLPC-----USERPPQHS 1004  
 QY 1051 D-----LRRAPQLNHSNDVVSNCNIRLVN-----QRIHFHLLGNL---WLSLKALYK 1099  
 Db 1005 DPLTOISRSPMLDCSTADCLQFCD---VPSFSVQBELDFTLKMLSPQWVR--ETLQKK 1059  
 QY 1100 SKIMV-NAALQKSPFIPREDSRQIEFISKQEDQVPTMIIVGSTLGGLLAL 1158  
 Db 1060 VUVSVAEITPOTSVISQLPGQAFMRQKQWVLEEDVINA-IFLIMSSVGAULLAL 1118  
 QY 1159 LVIALRKLGFRSARRRREPGLDTPK 1185  
 Db 1119 ITATLYLGLFF---KSHYKEMLEDKPE 1142

RESULT 10  
 ITAX HUMAN  
 ID ITAX HUMAN STANDARD; PRT; 1163 AA.  
 AC P20762;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Integrin alpha-X precursor (leukocyte adhesion glycoprotein p150,95  
 DE alpha chain) (leukocyte adhesion receptor p150,95) (CD11c) (leu M5).  
 GN ITGAX OR CD11C.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89166645; PubMed=3327687;  
 RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;  
 RT "cDNA cloning and complete primary structure of the alpha subunit of  
 RT a leukocyte adhesion glycoprotein, p150,95.";  
 RL EMBO J. 6:4023-4028(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90153906; PubMed=2303426;  
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;  
 RT "Genomic structure of an integrin alpha subunit, the leukocyte  
 RT p150,95 molecule.";  
 RL J. Biol. Chem. 265:2782-2788 (1990).  
 RN [3]  
 RP ERRATUM.  
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;  
 RL J. Biol. Chem. 265:12750-12751 (1990).  
 RN [4]  
 RP SEQUENCE OF 20-43.  
 RX MEDLINE=87167596; PubMed=3549901;  
 RA Miller L.J., Wiebe M., Springer T.A.;  
 RT "Purification and alpha subunit N-terminal sequences of human Mac-1  
 RT and p150,95 leukocyte adhesion proteins.";  
 RL J. Immunol. 138:2381-2383 (1987).  
 CC -!- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT  
 CC RECOGNIZES THE SEQUENCE G-P-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL  
 CC INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY  
 CC IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.  
 CC -!- SUBUNIT: HETEROIDIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X  
 CC ASSOCIATES WITH BETA-2.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND

CC CC GRANULOCYTES.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE. family.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD11c entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11c.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; M81695; AA59180.1; -;  
 CC EMBL; Y00093; CA68283.1; -;  
 CC EMBL; M29165; -; NOT ANNOTATED CDS.  
 CC EMBL; M29487; AA51620.1; ALT SEQ.  
 CC EMBL; M29482; AA51620.1; JOINED.  
 CC EMBL; M29483; AA51620.1; JOINED.  
 CC EMBL; M29484; AA51620.1; JOINED.  
 CC EMBL; M29485; AA51620.1; JOINED.  
 CC EMBL; M29486; AA51620.1; JOINED.  
 CC PIR; A36584; RWHDIC.  
 CC PDB; IN3Y; 18-FEB-03.  
 CC Genew; HGNC:6152; ITGAX.  
 CC MIM; 151510; -;  
 CC GO; GO:0008305; C:integrin complex; TAS.  
 CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.  
 CC GO; GO:0004872; F:receptor activity; TAS.  
 CC GO; GO:0007155; P:cell adhesion; TAS.  
 CC GO; GO:0007337; P:histogenesis and organogenesis; TAS.  
 CC InterPro; IPR000413; Integrin\_alpha.  
 CC InterPro; IPR002035; VWF\_A.  
 CC Pfam; PF01839; FG-GAP; 3\_  
 CC Pfam; PF00357; integrin\_A; 1.  
 CC Pfam; PF00092; vwa; 1.  
 CC PRINTS; PR01185; INTEGRINA.  
 CC PRINTS; PR0453; VWFADOMAIN.  
 CC SMART; SM00191; Int\_alpha; 5.  
 CC SMART; SMO0327; VWA; 1.  
 CC PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 CC PROSITE; PS50234; VWFA; 1.  
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 CC Signal; Magnesium; Calcium; Repeat; 3D-structure.  
 CC SIGNAL 1 19  
 FT CHAIN 20 1163 INTEGRIN ALPHA-X.  
 FT DOMAIN 20 1107 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1108 1128 POTENTIAL.  
 FT DOMAIN 1129 1163 CYTOPLASMIC (POTENTIAL).  
 FT REPEAT 34 87 FG-GAP 1.  
 FT REPEAT ? ? FG-GAP 2.  
 FT DOMAIN 165 351 VWFA.  
 FT REPEAT ? ? FG-GAP 3.  
 FT REPEAT 402 453 FG-GAP 4.  
 FT REPEAT 455 517 FG-GAP 5.  
 FT REPEAT 518 576 FG-GAP 6.  
 FT REPEAT 581 633 FG-GAP 7.  
 FT CA\_BIND 466 474 POTENTIAL.  
 FT CA\_BIND 530 538 POTENTIAL.  
 FT CA\_BIND 593 601 POTENTIAL.  
 FT SITE 1131 1135 GFFKR MOTIF.  
 FT DISULFID 69 76 BY SIMILARITY.  
 FT DISULFID 108 126 BY SIMILARITY.  
 FT DISULFID 655 712 BY SIMILARITY.  
 FT DISULFID 771 777 BY SIMILARITY.  
 FT DISULFID 848 863 BY SIMILARITY.  
 FT DISULFID 998 1022 BY SIMILARITY.  
 FT DISULFID 1027 1032 BY SIMILARITY.  
 FT CARBOHYD 61 61 N-LINKED (GLCNAC... ) (POTENTIAL).

|   |          |   |                               |             |
|---|----------|---|-------------------------------|-------------|
| FT  | CARBOHYD | 89  | N-LINKED (GLCNAC. . .)        | (POTENTIAL) |
| FT  | CARBOHYD | 392   | N-LINKED (GLCNAC. . .)        | (POTENTIAL) |
| FT  | CARBOHYD | 697   | N-LINKED (GLCNAC. . .)        | (POTENTIAL) |
| FT  | CARBOHYD | 735   | N-LINKED (GLCNAC. . .)        | (POTENTIAL) |
| FT  | CARBOHYD | 899   | N-LINKED (GLCNAC. . .)        | (POTENTIAL) |
| FT  | CARBOHYD | 939   | N-LINKED (GLCNAC. . .)        | (POTENTIAL) |
| FT  | CARBOHYD | 1050  | N-LINKED (GLCNAC. . .)        | (POTENTIAL) |
| FT  | CONFLICT | 490   | G -> A (IN REF. 2)            |             |
| FT  | CONFLICT | 756   | L -> D (IN REF. 2)            |             |
| SQ  | SEQUENCE | 1163 AA; 127885 MW; 6C4E19C3P62A473 CRC64;                    |                               |             |
| Query Match   |          |   |                               |             |
| Best Local Similarity 28.4%; Pred. No. 2.1e-65;                     |          |   |                               |             |
| Matches 354; Conservative 219; Mismatches 471; Indels 201; Gaps 48; |          |   |                               |             |
| QY  | 7        | LIVAWALSLWPGFTDFNMDTRKPRVPGSRAP                               | -----FGVTVQCHDISGNKVLVGA      | 60          |
| DB  | 10       | LFTALATSL   | -----GFNLDTSE                 | 53          |
| QY  | 61       | PLETNGYQKTDYKCPVHNGCTKLNLRVTLNYSERKDNRLGLSLATNPKDNSFLA        | 120                           |             |
| DB  | 54       | PQKITAANTGGLYQCYGTGACEPTGL                                    | -----QVPEAVNMSLGLSLASTTSPQLLA | 107         |
| QY  | 121      | CSPLWSHECGSSYITGMSRVNSNFRPSKTVAPALQRC-QTYMDIVIVLDSGSIYP       | --- 177                       |             |
| DB  | 108      | CGPTVHECEGRNNYLTGLCLLGT-QLTORLPVSRQECPRQEQDIVFLIDGSGSISRN     | 166                           |             |
| QY  | 178      | WVEQVFLINILKFKYIGPQQVGVQYCEDVHEFHLDNYSRVKDVVEASHIEQ-RG        | 236                           |             |
| DB  | 167      | FATMMNFVRAVISQFQ--RPSQFSLMQFSNKFQCHTFEERFRSTNPLSLLASVHQLQ     | 224                           |             |
| QY  | 237      | GTETRTAFGLIFARSAFOGKRGKAKKVMIVITGESH-DSPDLKVIQOQSERNVTRY      | 295                           |             |
| DB  | 225      | FTYATAIQVHRLPHASYGARRDATKILVITDGKGGSLDYKDVIPMDAAGIIRY         | 284                           |             |
| QY  | 296      | AVAV-LGYNNRGINPETFLNEIKYIASDPDDKHPNVDEALDKOIVDALGDRIFSLG      | 354                           |             |
| DB  | 285      | ALGVGLAFQNRNS-----WKLNDIASKPSQEHFKVEDPDAIKDIONQKEKIFAEG       | 338                           |             |
| QY  | 355      | T-NKNETSPGLEMSQTFGSHVVEDGVILGAVGAYDNGAVLKETSAKVIP--LRSEVL     | 411                           |             |
| DB  | 339      | TETTSSSSFELEMAQGFSAVTPDPGVILGAVGSPFTWSGAF-----LYPPNMSPTFI     | 391                           |             |
| QY  | 412      | KEPPEELKHGAYLVGTVTSVSSRQGRVYVAGAPENHTGVKILFTWHNNESLTHQAM      | 471                           |             |
| DB  | 392      | NMSQENVNDRSILGYS--TELALWKGVQSLVIGAPYQHTGKAVIPT--QVSQWRNKAEV   | 449                           |             |
| QY  | 472      | RQOQIGSYFGSEITSYDIDGDGVDVLLVGAPMYFNEGERGKYVYELRQ--NRFYVNG     | 529                           |             |
| DB  | 450      | TGTQIGSYFGASLCSVDVDTGSDTLVLICAPHYEQTR--GGQVSVCPPLRGWRWACDA    | 508                           |             |
| QY  | 530      | TKDSDSHYQNAFEGSSIASVRDLNODSYNDVVVGPALDNDHAGAYVIFHGFEG-SILKTP  | 588                           |             |
| DB  | 509      | VLFGQGHGHPWRFGAALTVLGDVNGDKLTDVWVIGAPGEBENRGAVLPHGVLPSPISPSH  | 569                           |             |
| QY  | 589      | KQRTASELATGLQYFGCSHIGOLDLNEEDGLIDLAVGALGNVILMSRPPVVOINASLHPE  | 648                           |             |
| DB  | 569      | SQRIAGSLSRLQYFGQALSGGQDLTDQGLDVLAVGARGQVLLKTRPVLVWGVSMQFI     | 628                           |             |
| QY  | 649      | PSKINIFHRDKES--GRDATCLAAFLCF-----TFIPLAPHTYTVGIRYNATMDRER     | 701                           |             |
| DB  | 629      | PASIPRAFECEQVVSQETLVOSNTCLYDKRSNRLGLSRLQSS-----VTLDLALDPGR    | 685                           |             |
| QY  | 702      | YTPRAHLDGEGDRFTNRVALLSGQBLCEIRINPHVL---DTADYKVPVTFVSVEYSLENDP | 757                           |             |
| DB  | 686      | LSPRATEQETKNSLSRVRVIGL-KAFCE--NFNLLPSCVEDSVTPITRLNFTLVGKP     | 742                           |             |
| QY  | 758      | -----DHGPMLDGWTTLRVSVFPWNGCNEDEHCVDPDLVDARSDLPATMEYCORVLK     | 812                           |             |
| DB  | 743      | LLAFRLRLMLAALAQRYFTASLPFXKCGADHIC-----                        | 777                           |             |
| QY  | 813      | PADCCSAYTSLFDTTTFIESTRQVAVATLENRGENAYSTVLNISOQANLQFASLI--     | 870                           |             |

|                                    |   |  |      |          |
|------------------------------------|---|--|------|----------|
| DB                                 | 778   | --QDNLGISFSPGLKSLLVGSNLLENAEVMVMDGSDSYGTTTTFSHPGAGLSRYVAEG | 835  |          |
| QY                                 | 871   | QKESDSDGIECVNEBRLOKO-----VCNVSYPFRAXAKVAFRLDSEFS-KSIFLHLE  | 924  |          |
| DB                                 | 836   | QKQQLSLSLHTCTDSAPVSGQSTWSTCRINHELLIFRGAQITFLATPVSRAVLGDRLL | 895  |          |
| QY                                 | 925   | IELAAGSDSNRSDGTKEKNVAPLRFHLKYEADVLFTRSSSLSHSYEVLNSS-----   | 975  |          |
| DB                                 | 896   | LTVNSSENNTPTSK--TTFQLELPVKY---AVTVVSSHEQFTKYLNFSESEKESHV   | 950  |          |
| QY                                 | 976   | LERYDGGIPPPSCIFRIQNLGL--FPIHGMMKTIPIATRSNGRLKLKDLFLTDEANT  | 1032 |          |
| DB                                 | 951   | AMHRY-----QVNNIGQDLFPV---SINFWPVE-----LNQEA--              | 981  |          |
| QY                                 | 1033  | SCNTWGN-----STEYRPTPVER---DLRRAPOLNHSNDVWSINCINRILVP       | 1076 |          |
| DB                                 | 982   | ---VMDVESHQPNQSLRCSSEKIAFPASDFLAHIQKPNVLDCSIAGCLAFRC       | 1035 |          |
| QY                                 | 1077  | N-----QEIFHLGNL---WLSLKALKYKSM---KIMVNAALQRFHSPIFREEDPSR   | 1126 |          |
| DB                                 | 1036  | SFSVQBEHLDFTLGNLSFGWVRQILQKQSVSVVABITFDTSVTSQLPQGFQAFRAQTT | 1095 |          |
| QY                                 | 1127  | QIB-PEISKQEDMQVPIWIVGSTLGGLLALLLALLVLRKLGPFR               | 1170 |          |
| DB                                 | 1096  | VLEIKYKHN-----PTPLVGSISGILLALLTAVLVKVGFFK                  | 1134 |          |
| RESULT 11                          |   |  |      |          |
| ITAM_MOUSE STANDARD; PRT; 1153 AA. |   |  |      |          |
| ID                                 | ITAM_MOUSE  | STANDARD;  | PRT; | 1153 AA. |
| AC                                 | P05555;   |  |      |          |
| DT                                 | 01-NOV-1988 (Rel. 09, Created)  |  |      |          |
| DT                                 | 01-FEB-1991 (Rel. 17, last sequence update)   |  |      |          |
| DT                                 | 21-FEB-2003 (Rel. 41, last annotation update)   |  |      |          |
| DE                                 | Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).  |  |      |          |
| GN                                 | ITGAM.  |  |      |          |
| OS                                 | Mus musculus (Mouse).   |  |      |          |
| OC                                 | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |  |      |          |
| OX                                 | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  |  |      |          |
| NCBI                               | TaxID=10090;  |  |      |          |
| NP                                 | SEQUENCE FROM N.A.  |  |      |          |
| RX                                 | MEDLINE=89312584; PubMed=3044779;   |  |      |          |
| RA                                 | Pyrela R.;  |  |      |          |
| RT                                 | "Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor."  |  |      |          |
| RL                                 | EMBO J. 7:1371-1378(1988).  |  |      |          |
| RN                                 | [2]   |  |      |          |
| RP                                 | SEQUENCE OF 11-45 FROM N.A.   |  |      |          |
| RC                                 | STRAIN=BALB/c; TISSUE=Spleen;   |  |      |          |
| EX                                 | MEDLINE=86287312; PubMed=2942940;   |  |      |          |
| RA                                 | Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E., Larson R.S., Roberts T.M., Springer T.A.;  |  |      |          |
| RT                                 | "A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1."   |  |      |          |
| RL                                 | Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).   |  |      |          |
| RN                                 | [3]   |  |      |          |
| RP                                 | SEQUENCE OF 17-28.  |  |      |          |
| RX                                 | MEDLINE=85188276; PubMed=3887182;   |  |      |          |
| RA                                 | Springer T.A., Teplow D.B., Dreyer W.J.;  |  |      |          |
| RT                                 | "Sequence homology of the IFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon."  |  |      |          |
| RL                                 | Nature 314:540-542(1985).   |  |      |          |
| CC                                 | !- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN |  |      |          |



Db 783 TMSAGLDT---LVVGGPQDFNMSVTLRNDGBDSYGTQVTVVYPSGLSYRKDSASQNP 839

QY 869 ----LQKEDSDSIECWNEERBLQKQVCNVSPPFRARAKVAFRLDSEF-SKSIPLHL 923

Db 840 KKPWFVKPABSSSSSE---GHGALKSTTWNINHPFPANSEVTFNFDVDSHSGNKL 896

QY 924 EIELAAGSDNERDSTKEDNVAPLRFHLKTYEADVLFRSSSLGHY-----EVLN 975

Db 897 LKAIIVASENMW---SRTHKTKFQLPVPKVIATVMTVSDSSIRYLNFTASEMTSKVIQH 954

QY 976 LEYDYGIP---PFCIP-----RIONGLPPIHGGMMKTIPIATRSGRLLKRLD 1028

Db 955 QYQFNELGQSLVSVVFPVQINNVTWD-HPQVI-----PSQ 993

QY 1029 EANTSNCNMGNSTEYRTP---VEEDLRAPQLHNSDVSINCLRLVPNOEI-NFHL 1084

Db 994 NLSSACH-----TEQSPHSNFRDQLERTFVLNCSVAVCKRIQCDLPSPNTOEII 1048

QY 1085 LGMH---WLRSK---ALKYKSKIMVNAALQRFSPFTFREEDPSRQIEFISQEDW 1138

Db 1049 KGNLSPDWYIKTSHGHLVSSTEILFN-----DSAFALLQGSYSYRSKTTKVB 1101

QY 1139 QV--PTWIIVSGTLGLLALLALVLRKLGFR 1170

Db 1102 EVHNPVLVSGSIGGLVLLALITAGLYKLGFEK 1135

RESULT 12

ITAM\_HUMAN STANDARD; PRT; 1152 AA.

AC P11215;

DT 01-JUL-1989 (Rel. 11, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MOI)

DE (Neutrophil adherence receptor).

GN ITGA OR CR3A OR CD11B.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RX MEDLINE=88315033; PubMed=2457584;

RA Corbi A.A., Kishimoto T.K., Miller L.J., Springer T.A.;

RT "The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";

RL J. Biol. Chem. 263:12403-12411(1988).

RN [2]

RX MEDLINE=88190151; PubMed=2833753;

RA Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;

RT "Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mo1: chromosomal localization and homology to the alpha subunits of integrins.";

RL Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).

RN [3]

RX MEDLINE=88257215; PubMed=2454931;

RA Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;

RT "Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mo1 (complement receptor type 3).";

RL J. Cell Biol. 106:2153-2158(1988).

RN [4]

RX MEDLINE=93123748; PubMed=8419480;

RA Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;

RT "Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";

RL J. Immunol. 150:480-490(1993).

RN [5]

RX MEDLINE=92073318; PubMed=1683702;

RA Shelley C.S., Arnaout M.A.;

RT "The promoter of the CD11b gene directs myeloid-specific and developmentally regulated expression.";

RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).

RN [7]

RX MEDLINE=92144986; PubMed=1346576;

RA Pahl H.L., Rosmarin A.G., Tenen D.G.;

RT "Characterization of the myeloid-specific CD11b promoter.";

RL Blood 79:865-870(1992).

RN [8]

RX MEDLINE=87076671; PubMed=3539202;

RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;

RT "N-terminal sequence of human leukocyte glycoprotein Mo1: conservation across species and homology to platelet IIb/IIIa.";

RL Biochim. Biophys. Acta 874:368-371(1986).

RN [9]

RX MEDLINE=95171458; PubMed=7867070;

RA Lee J.O., Rieu P., Arnaout M.A., Liddington R.;

RT "Crystal structure of the A domain from the alpha subunit of integrin CR3 (CD11b/CD18).";

RL Cell 80:631-638(1995).

RN [10]

RX MEDLINE=96363671; PubMed=8747460;

RA Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;

RT "Two conformations of the integrin A-domain (I-domain): a pathway for activation?";

RL Structure 3:1333-1340(1995).

RN [11]

RX MEDLINE=98362595; PubMed=9687375;

RA Balgwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A., Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrichson R.L., Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E., Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;

RT "Cation binding to the integrin CD11b I domain and activation model assessment.";

RL Structure 6:923-935(1998).

RN [12]

RX MEDLINE=98226734; PubMed=9560195;

RA Oxvig C., Springer T.A.;

RT "Experimental support for a beta-propeller domain in integrin alpha subunits and a calcium binding site on its lower surface.";

RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).

CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN.

CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M ASSOCIATES WITH BETA-2.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND







Db 781 ----SITFSFMSLCLVVGPPREFNFVTVTVRNDCGDSYRTQVTFPPDLISYRKVSTLQN 836  
Qy 873 EDSXGS--IECVNEER-----RLQKQVCNVSYPFRKAKAVARLDSEP-SKSIPLHLHL 924  
Db 837 QRQSRWRLACESASSTEVGALKSTCSINHIFPENSVTNITPVDVDSKASLGNKLL 896  
Qy 925 IELAAGSDSNEROSTKEDNVAPIRLFKYEADVLPTRSSLSHYEYKLNSSLERYDGI 984  
Db 897 LKANVTSENNPRNTKTE--FQLELPVKYAVYVMVTSHGVSSTKY--LNFTAS--ENVTSR 949  
Qy 985 PFCSCIFHQLGLFPIHGMWKITIPATSGNRLKLRDLFLDEANTSCNIWGNSTEYR 1044  
Db 950 VMQHQYQVSNLQ---RSLPISIVLPVFLVNTQVTDREQVTFSENLSSTC--HFKERL 1004  
Qy 1045 PTPVSE--EDLRRAPQLNHSNDVVSINCNIRLVP--NOEINPHLLGNL---WLSLX---A 1095  
Db 1005 PSHSDFLAELKAPVNVCSIAVCORIQCDIPFGIOBEFNATLKNLSFDPWYKTSNHL 1064  
Qy 1096 LKYSKMKIMNAALQROFHSPFFREDPSRQIE--EISKQEDNQVPMIIVGSTLQELL 1154  
Db 1065 LIVSTAEILFNSVFTLLPQGGAFVRSQTKTEKPEPEVN-----PLPDIVGSSVGGELL 1118  
Qy 1155 LLLALLVLALRLKGLFFRSARRRE-----PGLDP 1182  
Db 1119 LLALITAAVLKLGFEK--RQYKDMWSGGPPGABP 1151

RESULT 13  
ITAL MOUSE  
ID ITAL MOUSE STANDARD; RET; 1163 AA.  
AC P24063;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE integrin alpha-L precursor (leukocyte adhesion glycoprotein LFA-1  
DE alpha chain) (Leukocyte function associated molecule 1, alpha chain  
DE (CD11a).  
GN Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91268576; PubMed=2051027;  
RA Kaufmann Y., Tseng B., Springer T.A.;  
RT "Cloning of the murine lymphocyte function-associated molecule-1  
RT alpha-subunit and its expression in COS cells."  
RL J. Immunol. 147:363-374 (1991).  
RN [2]  
RP SEQUENCE OF 24-42.  
RX MEDLINE=85198276; PubMed=3887192;  
RA Springer T.A., Teplow D.B., Dreyer W.J.;  
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion  
RT glycoproteins and unexpected relation to leukocyte interferon."  
RL Nature 314:540-542 (1985).  
CC -!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,  
CC ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA  
CC INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL  
CC MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES  
CC AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L  
CC SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED  
CC LEUKOCYTES RECRUITMENT.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L  
CC ASSOCIATES WITH BETA-2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: LEUKOCYTES.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; N60778; AAA39426.1; .  
DR PIR; I56126; I56126.  
DR HSP; P20701; 1LEA.  
DR MGD; MGI:96506; Itgal.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01939; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PRO1135; INTEGRINA.  
DR PRINTS; PRO0453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; vwa; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS0234; VWFA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Signal; Magnesium; Calcium;  
KW Repeat.  
FT SIGNAL 1 23  
FT CHAIN 24 1163  
FT DOMAIN 24 1084  
FT TRANSMEM 1085 1108  
FT DOMAIN 1109 1163  
FT REPEAT 39 88  
FT REPEAT 2 7  
FT DOMAIN 148 334  
FT REPEAT 399 454  
FT REPEAT 455 514  
FT REPEAT 516 573  
FT REPEAT 576 628  
FT CA\_BIND 466 474  
FT CA\_BIND 528 536  
FT CA\_BIND 588 596  
FT SITE 1111 1115  
FT DISULFID 70 77  
FT DISULFID 108 126  
FT DISULFID 147 199  
FT DISULFID 651 705  
FT DISULFID 767 773  
FT DISULFID 840 856  
FT DISULFID 993 1009  
FT DISULFID 1017 1048  
FT CARBOHYD 86 86  
FT CARBOHYD 185 185  
FT CARBOHYD 270 270  
FT CARBOHYD 444 444  
FT CARBOHYD 668 668  
FT CARBOHYD 696 696  
FT CARBOHYD 724 724  
FT CARBOHYD 728 728  
FT CARBOHYD 776 776  
FT CARBOHYD 857 857  
FT CARBOHYD 880 880  
FT CARBOHYD 890 890  
FT CARBOHYD 899 899  
FT CARBOHYD 927 927  
FT CARBOHYD 1056 1056  
SQ SEQUENCE 1163 AA; 128343 MW; A7A3078469B8232F CRC64;  
Query Match 17.2%; Score 1073.5; DB 1; Length 1163;  
Best Local Similarity 28.1%; Pred. No. 14e-61;  
Matches 358; Conservative 206; Mismatches 450; Indels 261; Gaps 55;  
Qy 4 PRGLVAVNAL--SLWPGFTDFNMDTRKPRVPGSRTAFFGYVQQHDSGNKMLVWGP 61

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; N60778; AAA39426.1; .  
DR PIR; I56126; I56126.  
DR HSP; P20701; 1LEA.  
DR MGD; MGI:96506; Itgal.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01939; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PRO1135; INTEGRINA.  
DR PRINTS; PRO0453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; vwa; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS0234; VWFA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Signal; Magnesium; Calcium;  
KW Repeat.  
FT SIGNAL 1 23  
FT CHAIN 24 1163  
FT DOMAIN 24 1084  
FT TRANSMEM 1085 1108  
FT DOMAIN 1109 1163  
FT REPEAT 39 88  
FT REPEAT 2 7  
FT DOMAIN 148 334  
FT REPEAT 399 454  
FT REPEAT 455 514  
FT REPEAT 516 573  
FT REPEAT 576 628  
FT CA\_BIND 466 474  
FT CA\_BIND 528 536  
FT CA\_BIND 588 596  
FT SITE 1111 1115  
FT DISULFID 70 77  
FT DISULFID 108 126  
FT DISULFID 147 199  
FT DISULFID 651 705  
FT DISULFID 767 773  
FT DISULFID 840 856  
FT DISULFID 993 1009  
FT DISULFID 1017 1048  
FT CARBOHYD 86 86  
FT CARBOHYD 185 185  
FT CARBOHYD 270 270  
FT CARBOHYD 444 444  
FT CARBOHYD 668 668  
FT CARBOHYD 696 696  
FT CARBOHYD 724 724  
FT CARBOHYD 728 728  
FT CARBOHYD 776 776  
FT CARBOHYD 857 857  
FT CARBOHYD 880 880  
FT CARBOHYD 890 890  
FT CARBOHYD 899 899  
FT CARBOHYD 927 927  
FT CARBOHYD 1056 1056  
SQ SEQUENCE 1163 AA; 128343 MW; A7A3078469B8232F CRC64;  
Query Match 17.2%; Score 1073.5; DB 1; Length 1163;  
Best Local Similarity 28.1%; Pred. No. 14e-61;  
Matches 358; Conservative 206; Mismatches 450; Indels 261; Gaps 55;  
Qy 4 PRGLVAVNAL--SLWPGFTDFNMDTRKPRVPGSRTAFFGYVQQHDSGNKMLVWGP 61

Db 8 PRLLILGLQFAKAM-----SNDLDRPTQSFIAQGRHFGYQVLOIE-----DGVVVGAP 58  
Qy 62 LETNGYQKTDGVYKCPVIRHGNCTKMLNLRVTLSNVSERXDNMRRLGLSLATNPKDMSFLAC 121  
Db 59 GEGD---NTGGLYHCRTSSEFCOPVSLH-----GSNHTSKYLEMTLATDAAGSLLAC 108  
Qy 122 SPLWSEHCSSYVYTCMCVRNSNPFPSKTV-----APALQRC-QTYMDIVIVLDGSLN 173  
Db 109 DGLSRETCQNTYLSGLC-----YLFQSLBGPMLQNPAYQECMKGVLDVLFDFGSQ 162  
Qy 174 SI--YPMVBOVHPLINILKFKYIGPQIQGVVYQGEDVVFHFLNDRYSVK-----DVV 226  
Db 163 SLDRKDFEKILEFMKDMVK-----LSNTSYQFAAVQSPDCTCTETFLDY--VKQKNPQV 218  
Qy 227 BAASHIEQGGTE-----TRTAPGIEPARSEAP--QKGRKAKKVMIVITGESHDSPD 279  
Db 219 -----LGSVQPMFLANTFRAINVYVAVHFKESGARPDATKVLVITDEGEASDGN 270  
Qy 280 LEKVIQOSERDNYTRVAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFNNVDE--AAL 338  
Db 271 I-----SAADITRIIGIKHF--VSQKQKTLH--IFASEPVEE-FVKILTTFEKL 318  
Qy 339 KOIVDALGRIFSELETKNNE--TSFGLMSQTFSSHVVEDGVLLGAVGAYDWNGAVLKB 397  
Db 319 KOLFDTLQRIYAIEGTNRDLTSFNNELSSGISADLSKGEAVVAVGAKDW-----371  
Qy 398 TSAGKVIPIRE-----SYLKKEPEELKNHGAIVLYTWTSVVSRQGRVYVAGAPRNNHTG 452  
Db 372 --AGGEDLREDLQGAITFGQEPITSDVRGGYLGTVVAMWTRSRSPFLAAGAPRQHV 429  
Qy 453 KVLITFMHN-----NRSLSITHAMRQQQIGSFSGSEITSDVIDGDGVTDLVLVAGAPYFN 507  
Db 430 QVLLFQAPAGGRWNT-----QKIEGTQIGSYFGGELCSVDLDQDGEALLIGAPLPG 485  
Qy 508 EGBERKVVYELRQNFYNGTKDHSYQNRFGSSIASVRLDNDQSYNDVVVGCAPLE 567  
Db 486 EQR--GGRVYQRRQLFVMSLSQDGPVYGRFGAAITALTDTINGDRLTDVAVGAPLE 544  
Qy 568 DNHAGAIYFPHGRGSIKTPKQITASBLATGLQYFGCSHNGQLDNLNDEGLDIAVGL 627  
Db 545 EQ--GAVYIFNGKPGGLSPQSPQIOGAQVFPQIRWFGSIHGVGLDGLDLADVVGAE 602  
Qy 628 GNAVILWSPVVOINASLHFPBSKINIFHRDCKRSGRDATCTLAFL--CFTPIFLAPHQ 685  
Db 603 GRVWLSRPVVDVVTLSFSPSEIPEVHEVCSYSARBEQKHGKVLKACFRIKPLTPQ 662  
Qy 686 -TTVGIRYNATWDERRYTPRAHLDEGGDRETNRAVLLSSGOLCERINPHV-LDTADVV 743  
Db 663 GRLLANLSTQLDCHRMBSRGLFPDGSHELSTSI--TPKSLDLPFHFPICIQDLI 720  
Qy 744 KPVTFSEVYSLEDDHGMMLDDG-----WPT--TLRVSVPPWNGCNEDEHCVPLDVL 794  
Db 721 SPINVSINFTSLLEEGTPRDQGRAMQPILRPSIHTVTKIPEKNGCDKXCEANLJLS 780  
Qy 795 --ARSDLPAMEYQCVLRKPAQDCSAYTLPDTPVFIESTQRVAZEATLENENAY 852  
Db 781 SPARG-----PURLMSASL-----AVEWTLNSGEGD 810  
Qy 853 STVLNISQSANLQF--ASLIQ-----KEDSGSIECVAEERLQOV--CNVSYPPF 900  
Db 811 WVELDLDFPRLGSLFRXVEMQLPHSRMPVSCIELTEGS-----SLLTNKLKONVSSPIF 863  
Qy 901 RAKAKVAPR-----LDSEFSKIFLH--HLEIELAGSDSNERDSTKEDVAPLPHL 951  
Db 864 KAGQEVSLQVMFNTLLNSWEDFVGLNGTWCHC-----NENSSLOQEDNSAATHPV 914  
Qy 952 KYEADVL-----FT-----RSSLSLH--YEVKLNS-----LERYDGGIPPF 986  
Db 915 LYPVNLITKEQENSTLYISFTPKGPKTQVQVYQVRIQPSAYDNHNPFLALVGVPRH 974  
Qy 987 SCIFRQNLGLFPFHGMKMTIPIATRSGRLLKLRLDPLTDEANTSCNIGWNTSTYRPT 1046

Db 975 S-----EDLIT-----YTKSVQTDPLVT 992  
Qy 1047 PVEEDLRAPQLNHSNDVSVSINCNRILVNPNOEINFHLNML--RSLKALKYKSM--KI 1103  
Db 993 CHSEDLKR-PSSEABQPCLPQVPRCPVIFRWEILQVGTGVELSKETKASSTLSLCSL 1051  
Qy 1104 MVNAALQRPSPFIFREEDPSROIIBREISQEDMOPVPIWIIIVGTGSLGLLALLLVAL 1163  
Db 1052 SVFSNSGKHF--LYGSKASEAQLVKVDLIHEKM-LHVTIVLSGIGLGVLLFLFLAL 1107  
Qy 1164 RKLQFF-RSARRRE 1177  
Db 1108 YKVGFFKXNLKEME 1122  
RESULT 14  
ITAE MOUSE  
ID ITAE MOUSE STANDARD; PRT; 1167 AA.  
AC Q60677;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-E precursor (Integrin alpha M290).  
GN ITGAE.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AKR;  
RX MEDLINE=95187992; PubMed=7882170;  
RA Smith T.J., Ducharme L.A., Shaw S.K., Parker C.M., Brenner M.B.,  
RA Kileshaw P.J., Weis J.H.;  
RT "Murine M290 integrin expression modulated by mast cell activation.";  
RL Immunity 1:393-403(1994).  
CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT  
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL  
CC CELL MONOLAYERS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-E  
CC SUBUNIT GENE EXHIBIT A MARKED REDUCTION IN THE NUMBERS OF  
CC INTRAEPITHELIAL LYMPHOCYTES IN THE GUT AND IN THE DEVELOPMENT OF  
CC GUT-ASSOCIATED LYMPHOID AGGREGATES, SUPPORTING A SPECIFIC ROLE FOR  
CC THIS INTEGRIN IN MEDIATING RETENTION OF LYMPHOCYTES IN THE  
CC INTESTINAL WALL.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA  
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A  
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
CC EMBL; U12236; ARC52142.1; -.  
CC HSP; E1215; 1A8X.  
CC MGD; MGI:1298377; Itgae.  
CC InterPro; IPR000413; Integrin\_alpha.  
CC InterPro; IPR002035; VWFA.  
CC Pfam; PF01839; FG-GAP; 3.  
CC Pfam; PF00357; Integrin\_A; 1.  
CC Pfam; PF00092; vwa; 1.  
CC PRINTS; PR01185; INTEGRINA.  
CC PRINTS; PR00453; VWFADOMAIN.  
CC SMART; SM00191; Int\_alpha; 3.

DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00242; INTEGRIN ALPHA; 1.  
 DR PROSITE; PS0234; VWA; 1.  
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Signal; Repeat; Magnesium;  
 KW Calcium.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 1167 INTEGRIN ALPHA-E.  
 FT CHAIN 20 181 INTEGRIN ALPHA-E LIGHT CHAIN.  
 FT CHAIN 183 1167 INTEGRIN ALPHA-E HEAVY CHAIN.  
 FT DOMAIN 20 1114 EXTRACELLULAR (POTENTIAL).  
 FT TRANSEM 1115 1137 POTENTIAL.  
 FT DOMAIN 1138 1167 CYTOPLASMIC (POTENTIAL).  
 FT REPEAT ? ? FG-GAP 1.  
 FT REPEAT ? ? FG-GAP 2.  
 FT DOMAIN 149 192 X-DOMAIN (EXTRA DOMAIN).  
 FT DOMAIN 193 384 VWA.  
 FT REPEAT ? ? FG-GAP 3.  
 FT REPEAT 449 501 FG-GAP 4.  
 FT REPEAT 503 564 FG-GAP 5.  
 FT REPEAT 566 631 FG-GAP 6.  
 FT REPEAT 634 686 FG-GAP 7.  
 FT CA BIND 514 522 POTENTIAL.  
 FT CA BIND 578 586 POTENTIAL.  
 FT CA BIND 646 654 POTENTIAL.  
 FT DOMAIN 185 191 GLU-RICH (ACIDIC).  
 FT SITE 1140 1144 GPKR MOTIF.  
 FT DISULFID 72 83 BY SIMILARITY.  
 FT DISULFID 130 164 BY SIMILARITY.  
 FT DISULFID 698 754 BY SIMILARITY.  
 FT DISULFID 814 820 BY SIMILARITY.  
 FT DISULFID 898 898 BY SIMILARITY.  
 FT DISULFID 998 1023 BY SIMILARITY.  
 FT DISULFID 1031 1047 BY SIMILARITY.  
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 256 256 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 718 718 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 773 773 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 829 829 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 846 846 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 925 925 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 968 968 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1013 1013 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 1167 AA; 128983 MW; B8331C115DCCFFD CRC64;  
 Query Match 15.3%; Score 954; DB 1; Length 1167;  
 Best Local Similarity 26.2%; Pred. No. 8.2e-56;  
 Matches 340; Conservative 203; Mismatches 459; Indels 298; Gaps 51;  
 QY 10 AWALSMPGFTDFNNMDTRXPRVIGSRTPAGYVQOHDISGNKWLWVGCAPLETNGYOK 69  
 DB 27 AWTALQPG-----APAVLS-----SLLHLDPSNNQCLLVARRSSN--RN 65  
 QY 70 TGDVYK-----PVIGNCTKLNGLRVTLSNVSRKDNRLGLSLATWPKNS 117  
 DB 66 TAALRYCAISIDPEACQVEHCMPK---GRYQ-----GVTLVGN--HNG 107  
 QY 118 FLACSPWSECGS-SYITGMCSRVNSNF----- 146  
 DB 108 VLVCIQVQARKFRSLNSELNGACSLTLPNLDLQAQVAFSDLEGFLDPAHVDSDGYCRSK 167  
 QY 147 -----RSKTVAPALQRCQYMDIVVLDGNSNITYP--WVEVOHFLINILKQY 193  
 DB 168 GSGTGEKKSGARRRTVEEDBE-EDGTEIAIVLDGSGSIGPSDFQKAKNFISTMWRNFY 226

QY 194 IGPQIQVGVVOYGEDVVEHFLNDYRSVKDVVEAASHIEHQGGTGTETRTAPGIEFARS 253  
 DB 227 EKCFCENFALVOYGAVIQTEFDLQESRDINASLAKVQSIQVQKEV-TKTASAMQHVLDMI 285  
 QY 254 F--QXGGRKGAKKMIVITDGESHSP-DLEKVIQOESRDNUTRVAVALGYNNRGINP 310  
 DB 286 FIPSGSRKALKVMVLTGDIQFDPLNLTIVINSPRQGVVFAIGV---GDRFKN 341  
 QY 311 ETLFLNEIKYIASDPDDKHPFNVTDBAALKDIIDALGDRIFSLGEGTNKNETSFCLEMSQTS 370  
 DB 342 NTY-RELKLIASDPKEAHTPKYNTYSALDGLLSKLQORIVHMEGTVGDAQY--QLAQITG 398  
 QY 371 FSHHVDEG-VLLGAVGAYDN-GAVLKETSAGKVIPLRESYLKFPPELKN--HGAYLG 426  
 DB 399 FSAQILDRGQVLLGTGVAFFNMSGGALLYSTONG-----RGCFLNQTAKEDESTRVQSYLG 453  
 QY 427 YTVTSVSSRQGRVYVAGAPFNHTGKVLFTMHNNRSILITHOAMRGQOIGSYVGFSEITS 486  
 DB 454 YSL-AVLHKAHGISYVAGAPRHLKLGAVFELKXEDREDAFVRRIEGQMGSYFSGVLC 512  
 QY 487 VDIIDGVTVDLLVGAPMYFNREGRGKVVYVELRON--RFVYNGTLDKDSHYQNAFPGS 544  
 DB 513 VDIIDMGITDFELVAAPFYHIRG-BEGRVYVYVQVPEQDASFSLAHTLSGHPGLTNSRFG 571  
 QY 545 SIASVRDLNODSYNDVVVGCAPLEDNHA-----GAIYIFHGPRGSILKTPKORITASELA 598  
 DB 572 ANAAVGDINQAKFTDVAIGAPLEGAGDGASGYGVYIYNGHSGHLYDSPSOIARASVA 631  
 QY 599 TGLQVFGCSIHQDLINBDGLIDLAVGALGNAILMSPVQVAINASLHPEPS----- 650  
 DB 632 SGLHYFGMSVSGGLDPNGDGLADITVGSRDSAVLRSRPVVDLTVSMFTPDALPMVFTG 691  
 QY 651 --KINI-FHRDCKRSRDRATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYTPRAH 707  
 DB 692 KMDVNLCEVDSSVVASPEGLREMFNFT-----VDVDTKORQLQ 733  
 QY 708 LDE-----CGDRFTNRAVLSSQGLCELCERINFHVLDTADYVXPTVFSVSLSD 756  
 DB 734 CBDSSGQCLRKMGSGFLCEHFWL1STEELCE-----DCFSNITIKTVYEFOT 784  
 QY 757 P----DH-GPMLDDGWPITLVSVPFWNGCNEDEHCVPLVDLSDLPAMEYQORVLR 811  
 DB 785 SGRDRYDNPNTLDHYKEPSAIPOLPYEKDKNKVFCIAEQ----- 826  
 QY 812 KPAQDCSAYTLSDFTTVFIIESTRQVAVEATLENRGENAYSTVLNISOANLQASLIQ 871  
 DB 827 -----TTWISQQLVGVGTKE-VTMNLSLTNSGDSYTMNMAVPRNLQPKK-IQ 875  
 QY 872 KEDSGSECVNEBERRLOQV--QNVSPFPRAKAKVAPRLDSEFSKSIPLHLELAA 929  
 DB 876 REVSP-DVQCDPDPKFAVSLVMNCKIGHPIPK-RSSVNVSVTWQLEESVFPNRTADITVT 933  
 QY 930 QSDSNERDSTEDNVAPLRPHLKVEA-----DVLFP--TRSSLSHLYVKLSSLSRYDGI 982  
 DB 934 ISNSNEKSLAB--TRSLQFPHAFIAVLRSRPSVMYMTSQSPSHKEFPFENVHGENL--- 988  
 QY 983 GPPFCIFRIQNLGFLPIHGMMKTIPIATRSNGRLKLDRDFTDEANTSCNIWGNSTE 1042  
 DB 989 ---FGAVFQLQ-----ICVPIQLQD-FQIVRVKNTLKTQDTEC----- 1023  
 QY 1043 YRPTVEEDLRAPOLNH-----SNSDVVSINCMRLVPMOEIFHLLGNLWL 1090  
 DB 1024 ---TOSQBPACSDPVQVKEHWSVCAITNKENVTVAELSVGHTKQL-----L 1071  
 QY 1091 RSLKALKY-----KSMKIMVNAALQOFHSPFFREEDPSRQIEPEISKQEDQVPIW 1143  
 DB 1072 RDVSELPIGBISFNKSLYEGINAENHRT-KITVIFLKEBETRSILP----- 1116  
 QY 1144 IIVGSLGGLALLALLVLLALRKLGFRRSARRRRRRLDPT 1183  
 DB 1117 LIIGSISGLLVVWVITAILFKCGFFK--RKYQQLNEST 1154

```

RESULT 15
ID ITAE HUMAN STANDARD; PRT; 1179 AA.
AC P38570: Q9NZU9.
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1
DE antigen) (CD103 antigen) (Integrin alpha-TEL).
GN ITGAE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN [2]
RN [3]
RP REVISIONS TO 88-114.
RA Parker C.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 53-1179 FROM N.A.
RC TISSUE=Fetal kidney; PubMed=10673275;
RX MEDLINE=20138496; PubMed=10673275;
RA Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G.,
RA Stotlersuk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A.,
RA Green E.D.;
RT "The genomic region encompassing the nephropathic cystinosis gene
RT (CTNG): complete sequencing of a 200-kb segment and discovery of a
RT novel gene within the common cystinosis-causing deletion.";
RL Genome Res. 10:165-173(2000).
RN [4]
RP MUTAGENESIS OF ASP-109 AND PHE-316.
RX MEDLINE=20400502; PubMed=10837471;
RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y.,
RA Brenner M.B.;
RT "The role of alpha and beta chains in ligand recognition by beta 7
RT integrins.";
RL J. Biol. Chem. 275:25652-25664(2000).
CC -1- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN
CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL
CC EPITHELIAL CELLS.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD103 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd103.htm".
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL; L25851; AAB59359.2; --

```

```

DR EMBL; AF168787; AAF43107.1; --
DR PIR; A53213; A53213.
DR HSSP; P11215; LABX.
DR Genew; HGNC:6147; ITGAE.
DR MiM; 604682; --
DR GO; GO:0008305; C: integrin complex; TAS.
DR GO; GO:0004895; F: cell adhesion receptor activity; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 3.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Polymorphism; Magnesium; Calcium.
FT CHAIN 1 18 INTEGRIN ALPHA-E.
FT CHAIN 19 1179 INTEGRIN ALPHA-E LIGHT CHAIN.
FT CHAIN 179 1177 INTEGRIN ALPHA-E HEAVY CHAIN.
FT DOMAIN 19 1124 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1125 1147 POTENTIAL.
FT DOMAIN 1148 1179 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 181 198 GLU-RICH (ACIDIC).
FT REPEAT ? ? FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT REPEAT ? ? X-DOMAIN (EXTRA DOMAIN).
FT REPEAT ? ? VWFA.
FT REPEAT 401 456 FG-GAP 3.
FT REPEAT 457 506 FG-GAP 4.
FT REPEAT 510 571 FG-GAP 5.
FT REPEAT 573 638 FG-GAP 6.
FT REPEAT 641 693 FG-GAP 7.
FT CA_BIND 522 530 POTENTIAL.
FT CA_BIND 586 594 POTENTIAL.
FT CA_BIND 654 662 POTENTIAL.
FT SITE 1150 1154 GFFKR MOTIF.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 126 159 BY SIMILARITY.
FT DISULFID 706 762 BY SIMILARITY.
FT DISULFID 823 829 BY SIMILARITY.
FT DISULFID 833 907 BY SIMILARITY.
FT DISULFID 1008 1033 BY SIMILARITY.
FT DISULFID 1041 1057 BY SIMILARITY.
FT CARBOHYD 49 49 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 321 321 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 726 726 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 782 782 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 857 857 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 934 934 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 954 954 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1065 1065 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1096 1096 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 360 360 D -> E.
FT VARIANT 1041 1041 /FTID=VAR_008884.
FT VARIANT 1041 1041 C -> S.
FT MUTAGEN 208 208 D->A: LOSS OF E-CADHERIN BINDING.
FT MUTAGEN 316 316 F->A: LOSS OF E-CADHERIN BINDING.
FT CONFLICT 477 477 V -> I (IN REF. 3).
FT CONFLICT 482 482 Q -> R (IN REF. 3).
FT CONFLICT 950 950 R -> W (IN REF. 3).
FT CONFLICT 1019 1019 A -> V (IN REF. 3).
SQ SEQUENCE 1179 AA; 130088 MW; E558902EDF9D95E1 CRC64;
Query Match 14.9%; Score 924.5; DB 1; Length 1179;
Best Local Similarity 26.3%; Pred. No. 6.9e-52;

```

| Matches | 338; | Conservative | 206; | Mismatches | 482;   | Indels | 259;  | Gaps | 52;   |      |        |      |     |     |
|---------|------|--------------|------|------------|--------|--------|-------|------|-------|------|--------|------|-----|-----|
| Qy      | 23   | FNMTRKPRVPI  | PGSR | TAF        | GVTVQO | HDIS   | GNK-- | KLVV | GAPLE | TNGY | QKTDGV | KCPV | IHG | 81  |
| Db      | 19   | FNDV         | AR   | PWLT       | PKG    | AP     | FLV   | SL   | LLH   | QD   | PN     | TQ   | TL  | 75  |
| Qy      | 82   | NCT          | KL   | NR         | VT     | LS     | NV    | S    | R     | K    | N      | R    | L   | 138 |
| Db      | 76   | EIL          | CH   | P          | VE     | H      | P     | I    | P     | K    | R      | H    | R   | 125 |
| Qy      | 139  | CS           |      |            |        |        |       |      |       |      |        |      |     | 159 |
| Db      | 126  | CS           | L    | G          | P      | D      | L     | R    | P     | Q    | A      | N    | F   | 185 |
| Qy      | 160  |              |      |            |        |        |       |      |       |      |        |      |     | 205 |
| Db      | 186  | ED           | K    | E          | E      | E      | D     | E    | B     | E    | E      | R    | A   | 245 |
| Qy      | 206  | Y            | G    | E          | D      | V      | H     | E    | H     | L    | N      | D    | R   | 263 |
| Db      | 246  | Y            | G    | G          | I      | O      | T     | E    | F     | L    | R      | S    | O   | 304 |
| Qy      | 264  | K            | M    | I          | V      | I      | D     | G    | E     | S    | H      | D    | S   | 317 |
| Db      | 305  | K            | M    | V          | L      | T      | D     | G    | G     | I    | F      | E    | D   | 354 |
| Qy      | 318  | K            | Y    | T          | A      | S      | D     | P    | D     | K    | H      | F    | N   | 377 |
| Db      | 355  | N            | L    | I          | A      | S      | D     | P    | E     | T    | A      | F    | K   | 412 |
| Qy      | 378  | D            | -    | G          | V      | L      | L     | G    | A     | V    | A      | D    | N   | 435 |
| Db      | 413  | E            | R    | V          | L      | L      | G     | A    | V     | A    | D      | N    | 470 |     |
| Qy      | 436  | R            | Q    | E          | V      | T      | A     | G    | A     | P    | R      | N    | H   | 495 |
| Db      | 471  | T            | C    | S          | L      | S      | V     | Y    | A     | G    | A      | P    | Y   | 529 |
| Qy      | 496  | D            | V    | L          | V      | A      | G     | M    | P     | F    | N      | E    | G   | 553 |
| Db      | 530  | D            | P    | L          | L      | V      | A     | P    | F     | V    | H      | V    | G   | 588 |
| Qy      | 554  | Q            | S    | D          | S      | N      | D     | V    | V     | G    | A      | P    | L   | 607 |
| Db      | 589  | Q            | D    | K          | L      | T      | D     | V    | A     | I    | G      | A    | P   | 648 |
| Qy      | 608  | I            | H    | G          | O      | L      | D     | N    | E     | D    | G      | L    | D   | 667 |
| Db      | 649  | M            | A    | G          | G      | F      | D     | I    | S     | G    | L      | A    | D   | 699 |
| Qy      | 668  | C            | L    | A          | A      | P      | C     | F    | T     | I    | F      | L    | A   | 710 |
| Db      | 700  | V            | N    | V          | R      | L      | C     | F    | E     | --   | I      | S    | S   | 756 |
| Qy      | 711  | G            | G    | R          | F      | T      | N     | R    | A     | V    | L      | L    | S   | 765 |
| Db      | 757  | S            | G    | S          | O      | L      | C     | E    | D     | I    | L      | L    | M   | 807 |
| Qy      | 766  | G            | W    | P          | T      | L      | R     | V    | S     | V    | P      | P    | N   | 825 |
| Db      | 808  | Y            | T    | P          | P      | A      | I     | F    | O     | P    | Y      | E    | K   | 842 |
| Qy      | 826  | T            | T    | V          | I      | I      | E     | S    | T     | R    | V      | A    | E   | 885 |
| Db      | 843  | E            | --   | L          | V      | L      | G     | T    | K     | E    | L      | T    | N   | 897 |
| Qy      | 886  | R            | L    | Q          | K      | O      | V     | --   | C     | N    | V      | S    | P   | 940 |
| Db      | 898  | P            | V    | A          | S      | V      | I     | M    | C     | R    | I      | G    | H   | 953 |
| Qy      | 941  | E            | D    | N          | V      | A      | P     | L    | R     | F    | L      | K    | Y   | 994 |
| Db      | 954  | N            | E    | --         | T      | H      | T     | Q    | F     | R    | G      | F    | V   | 998 |

Search completed: June 24, 2004, 17:56:10  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 17:53:06 ; Search time 57 Seconds  
(without alignments)  
6576.064 Million cell updates/sec

Title: US-09-980-403-2  
Perfect score: 6224  
Sequence: 1 MDPRGLVVAWLSLWPGFT.....PRSSRRRREPCLDPTPKVLE 1188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL\_25:.\*  
1: sp\_archaea:.\*  
2: sp\_bacteria:.\*  
3: sp\_fungi:.\*  
4: sp\_human:.\*  
5: sp\_invertebrate:.\*  
6: sp\_mammal:.\*  
7: sp\_mhc:.\*  
8: sp\_organelle:.\*  
9: sp\_phase:.\*  
10: sp\_plant:.\*  
11: sp\_ricent:.\*  
12: sp\_virus:.\*  
13: sp\_vertebrate:.\*  
14: sp\_unclassified:.\*  
15: sp\_rvirus:.\*  
16: sp\_bacteriap:.\*  
17: sp\_archaeap:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID     | Description         |
|------------|--------|-------------|--------|-----------|---------------------|
| 1          | 5690   | 91.4        | 1188   | 11 Q7TQC3 | Q7TQC3 mus musculus |
| 2          | 4272   | 68.6        | 823    | 4 Q8WY18  | Q8WY18 homo sapien  |
| 3          | 3840   | 61.7        | 823    | 11 Q8CE84 | Q8CE84 mus musculus |
| 4          | 2135.5 | 34.3        | 1171   | 13 Q42094 | Q42094 gallus galli |
| 5          | 1196   | 19.2        | 1161   | 11 Q8QVE7 | Q8QVE7 rattus norv  |
| 6          | 1160   | 18.2        | 1169   | 11 Q9QXH4 | Q9QXH4 mus musculus |
| 7          | 1131.5 | 18.2        | 1169   | 4 Q8IV46  | Q8IV46 homo sapien  |
| 8          | 1125.5 | 18.1        | 1196   | 13 Q98TF1 | Q98TF1 cyprinus ca  |
| 9          | 1096.5 | 17.6        | 1151   | 11 Q9J130 | Q9J130 rattus norv  |
| 10         | 1065.5 | 17.1        | 1161   | 11 Q9WTV4 | Q9WTV4 mus musculus |
| 11         | 1063   | 17.1        | 1160   | 11 Q9R200 | Q9R200 mus musculus |
| 12         | 1045   | 16.8        | 1187   | 13 Q98TF0 | Q98TF0 cyprinus ca  |
| 13         | 1040   | 16.7        | 780    | 13 Q06271 | Q06271 xenopus lae  |
| 14         | 1018.5 | 16.4        | 1086   | 4 Q96H81  | Q96H81 homo sapien  |
| 15         | 979    | 15.7        | 1167   | 11 Q88341 | Q88341 rattus norv  |
| 16         | 954    | 15.3        | 1167   | 11 Q88340 | Q88340 rattus norv  |

17 904.5 14.5 927 6 Q8HZV0  
18 900 14.5 1160 6 Q8MKP4  
19 895 14.4 920 6 Q28984  
20 891 14.3 1038 11 Q8BS01  
21 881 14.2 1332 5 Q9BP08  
22 856 13.8 895 11 Q9WTF8  
23 848.5 13.6 1036 11 Q8CA73  
24 795.5 12.8 1036 11 Q91YD5  
25 742 11.9 1041 5 Q9UB90  
26 738 11.9 1041 5 Q76378  
27 719.5 11.6 1054 5 Q9UGS1  
28 691.5 11.1 1032 11 Q61989  
29 691.5 11.1 1033 6 Q9BGU3  
30 667.5 10.7 323 11 Q8CB84  
31 630 10.1 1119 5 Q86888  
32 604.5 9.7 1034 13 Q98TT7  
33 597 9.6 632 4 Q8NGH6  
34 595.5 9.6 1053 11 Q80YPS  
35 587 9.4 1132 11 Q80218  
36 584 9.4 1033 13 Q42598  
37 582.5 9.4 1073 11 Q8CC06  
38 571.5 9.2 1474 5 Q86887  
39 567.5 9.1 1016 13 Q91779  
40 556 8.9 604 11 Q8BQ25  
41 551.5 8.9 974 11 Q924W2  
42 547 8.8 1049 5 Q8SY51  
43 545 8.8 1047 6 Q9MZD6  
44 544.5 8.7 1007 6 Q9GK48  
45 539.5 8.7 833 5 Q9BFQ7

#### ALIGNMENTS

#### RESULT 1

Q7TQC3 ID Q7TQC3 PRELIMINARY; PRT; 1188 AA.  
AC Q7TQC3;  
DT 01-OCT-2003 (T-EMBLrel. 25, Created)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
DE All integrin.  
GN ITGAL1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Tiger C.-F., Popova S.N., Velling T., Ekman S., Forsberg E.,  
RA Gullberg D.;  
RT "allbi integrin is important for mesenchymal cell function:  
RT elimination of allbi leads to dwarfism."  
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Johansson M., Popova S.N.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
OC EMBL; AY124460; AAM62130.1; -;  
KW Integrin.  
SQ SEQUENCE 1188 AA; 133067 MW; 69B2BF20D75E4271 CRC64;

Query Match 91.4%; Score 5690; DB 11; Length 1188;  
Best Local Similarity 90.2%; Pred. No. 0;  
Matches 1071; Conservative 57; Mismatches 60; Indels 0; Gaps 0;  
Qy 1 MDPRGLVVAWLSLWPGFTDTFNMDTRKPRVIFPGSKTAPFGYTVQOHDISGNKWLVYGA 60  
Db 1 MDPRGLVVAWLSLWPGFTDTFNMDTRNPRVIAGPSAAFFGYTVQOHDISGNKWLVYGA 60  
Qy 61 PLFTNGYQKTDYKCPVHGNCTKLNGLRVTLNVSEKDNMRGLGLSLATNPKNDSFLA 120  
Db 61 PMETNGHQKTDYKCPVHGNCTKLNGLRVTLNVSEKDNMRGLGLSLATNPKNDSFLA 120

Qy 121 CSPLMSHECCSSYTTTCMSRPNNSFRFSKTVAPALQRCOTYNDIIVILDGNSIYPWVE 180  
 Db 121 CSPLMSHECCSSYTTTCMSRPNNSFRFSKTVAPALQRCOTYNDIIVILDGNSIYPWVE 180  
 Qy 181 VOHFLNILKFFIIGPQIQGVVQYCEVDVHFHFLNDYRSVXDVVAAASHIEQRGSTET 240  
 Db 181 VOHFLNILKFFIIGPQIQGVVQYCEVDVHFHFLNDYRSVXDVVAAASHIEQRGSTET 240  
 Qy 241 RTAFGEFARSEAFQGGKAGKVMIVITDGHSDSPOLEKVIQOSERDNTYRVAVL 300  
 Db 241 RTAFGEFARSEAFQGGKAGKVMIVITDGHSDSPOLEKVIQOSERDNTYRVAVL 300  
 Qy 301 GYNNRGINPETFLNEIKYIASDDKHGFNVVDEAALKDIDVADGDRIFSLEGTKNET 360  
 Db 301 GYNNRGINPETFLNEIKYIASDDKHGFNVVDEAALKDIDVADGDRIFSLEGTKNET 360  
 Qy 361 SFGLEMSQTFSSHVEDGVLGAVGAYDNGAVLTKTSAGKVIPLRESYLKFPPELKN 420  
 Db 361 SFGLEMSQTFSSHVEDGVLGAVGAYDNGAVLTKTSAGKVIPLRESYLKFPPELKN 420  
 Qy 421 HGAYLGYTTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNNRSLTHQAMRQOQIGSYF 480  
 Db 421 HGAYLGYTTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNNRSLTHQAMRQOQIGSYF 480  
 Qy 481 GSEITSVDIDGQVTDVLLVGAPMYPNEGRERKQVYVYELQNRQFVYNGTLKDSHSYQNA 540  
 Db 481 GSEITSVDIDGQVTDVLLVGAPMYPNEGRERKQVYVYELQNRQFVYNGTLKDSHSYQNA 540  
 Qy 541 RFGSSIASVRDLNQDSNDVWVGAPLEDNHAGAIYIFHGFPGSILKTPKORITASELATG 600  
 Db 541 RFGSSIASVRDLNQDSNDVWVGAPLEDNHAGAIYIFHGFPGSILKTPKORITASELATG 600  
 Qy 601 LQVFGCSIHQOLDNEDGLDLAGVAGNAILWSRPVVOINASLHFPESKINIFHRDCK 660  
 Db 601 LQVFGCSIHQOLDNEDGLDLAGVAGNAILWSRPVVOINASLHFPESKINIFHRDCK 660  
 Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYTPRAHLDGGRFTNRAV 720  
 Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYTPRAHLDGGRFTNRAV 720  
 Qy 721 LLSGGQELCERINPHVLDATYVKPVTFVSVEYSLEDPDHGPMDDGMPITLRSVFPFWNG 780  
 Db 721 LLSGGQELCERINPHVLDATYVKPVTFVSVEYSLEDPDHGPMDDGMPITLRSVFPFWNG 780  
 Qy 781 CNEDERHCVPLVLDARSDLTAMEYCORVLKPAQDCSAYTLSPDTTTFVITESTRORAV 840  
 Db 781 CNEDERHCVPLVLDARSDLTAMEYCORVLKPAQDCSAYTLSPDTTTFVITESTRORAV 840  
 Qy 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKOVNVSYPFF 900  
 Db 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKOVNVSYPFF 900  
 Qy 901 RAKAKVAFRLDSFPSKIFLHLEIELAAGSDSNERDSTKEDNVAPLPHLKYEADVLPT 960  
 Db 901 RAKAKVAFRLDSFPSKIFLHLEIELAAGSDSNERDSTKEDNVAPLPHLKYEADVLPT 960  
 Qy 961 RSSLSHYEVLKNSLERYDGIPTPSCIFRIONLGLFPIHGMWMTTPIATRSNRL 1020  
 Db 961 RSSLSHYEVLKNSLERYDGIPTPSCIFRIONLGLFPIHGMWMTTPIATRSNRL 1020  
 Qy 1021 KLADFLEANTSCNIGNSTERYPTVEEDLRRAPOLNHSNDVVSINCLRNPNQEI 1080  
 Db 1021 KLADFLEANTSCNIGNSTERYPTVEEDLRRAPOLNHSNDVVSINCLRNPNQEI 1080  
 Qy 1081 NEHLGLNLWLSLKALYKSMKIMVNAALQORHSPFIFREEDPSQIPEIISKQEDWQV 1140  
 Db 1081 NEHLGLNLWLSLKALYKSMKIMVNAALQORHSPFIFREEDPSQIPEIISKQEDWQV 1140  
 Qy 1141 PIWIVGSTLGGLLALLVALLKLGFFRSARRRPPGLDTPKVL 1188  
 Db 1141 PIWIVGSTLGGLLALLVALLKLGFFRSARRRPPGLDTPKVL 1188

## RESULT 2

Q8WYI8 PRELIMINARY; PRT; 823 AA.  
 ID Q8WYI8;  
 AC Q8WYI8;  
 DT 01-MAR-2002 (TREMELrel. 20, Created)  
 DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
 DE MSTP018.  
 GN MST018.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Aorta;  
 RA Xu Y.Y., Sun L.Z., Wu Q.Y., Liu Y.Q., Liu B., Zhao B., Wang X.Y.,  
 RA Song L., Ye J., Sheng H., Gao Y., Zhang C.L., Zhang J., Wei Y.J.,  
 RA Sun Y.H., Jiang Y.X., Zhao X.W., Liu S., Liu L.S., Ding J.P.,  
 RA Gao R.L., Qiang B.Q., Yuan J.G., Liew C.C., Zhao M.S., Rui R.T.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF111799; AAL39001.1;  
 DR GO; GO:0008305; C:integrin complex; IEA.  
 DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.  
 DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR PRINTS; PRO1185; INTEGRINA.  
 DR SMART; SM00191; Int alpha; 4.  
 SQ SEQUENCE 823 AA; 92672 MW; DE4E78079DCD4925 CRC64;

Query Match 68.6%; Score 4272; DB 4; length 823;  
 Best Local Similarity 99.6%; Pred. No. 2.9e-288;  
 Matches 820; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 366 MSOTCFSSHVEDGVLGAVGAYDNGAVLTKTSAGKVIPLRESYLKFPPELKNHAYL 425  
 Db 1 MSOTCFSSHVEDGVLGAVGAYDNGAVLTKTSAGKVIPLRESYLKFPPELKNHAYL 60  
 Qy 426 GYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNNRSLTHQAMRQOQIGSYFSGEIT 485  
 Db 61 GYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNNRSLTHQAMRQOQIGSYFSGEIT 120  
 Qy 486 SYVDIDGQVTDVLLVGAPMYPNEGRERKQVYVYELQNRQFVYNGTLKDSHSYONARFGSS 545  
 Db 121 SYVDIDGQVTDVLLVGAPMYPNEGRERKQVYVYELQNRQFVYNGTLKDSHSYONARFGSS 180  
 Qy 546 TASVRDLNQDSNDVWVGAPLEDNHAGAIYIFHGFPGSILKTPKORITASELATGLQYFG 605  
 Db 181 TASVRDLNQDSNDVWVGAPLEDNHAGAIYIFHGFPGSILKTPKORITASELATGLQYFG 240  
 Qy 606 CSIHGOLDNEDGLDLAGVAGNAILWSRPVVOINASLHFPESKINIFHRDCKRSGRD 665  
 Db 241 CSIHGOLDNEDGLDLAGVAGNAILWSRPVVOINASLHFPESKINIFHRDCKRSGRD 300  
 Qy 666 ATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYTPRAHLDGGRFTNRAVLLSSG 725  
 Db 301 ATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYTPRAHLDGGRFTNRAVLLSSG 360  
 Qy 726 QELCERINPHVLDATYVKPVTFVSVEYSLEDPDHGPMDDGMPITLRSVFPFWNGCNEDE 785  
 Db 361 QELCERINPHVLDATYVKPVTFVSVEYSLEDPDHGPMDDGMPITLRSVFPFWNGCNEDE 420  
 Qy 786 HCVPLVLDARSDLTAMEYCORVLKPAQDCSAYTLSPDTTTFVITESTRORVAVEATLE 845  
 Db 421 HCVPLVLDARSDLTAMEYCORVLKPAQDCSAYTLSPDTTTFVITESTRORVAVEATLE 480  
 Qy 846 NRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKOVNVSYPFPRAKAK 905  
 Db 481 NRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKOVNVSYPFPRAKAK 540  
 Qy 906 VAFRLDSFPSKIFLHLEIELAAGSDSNERDSTKEDNVAPLPHLKYEADVLFRSSSL 965







[illegible]

```

Db      819  ---SPFIVSRNDRFTIQISIKNKKDSAVNTRALVOYSPNIIFAGIEDIOKDCSES--- 872
Qy      882  VNEERLRQKQCVNVPFFFRAKAVFRIDSEFSKGFIFLHHLBIELAAGSDSNERDSTKE 941
Db      873  -----HNITCKVGPFLKPEEBSFKISQFNASVLLLENATVDVVATSDSEPPETLR 925
Qy      942  DNVAPLRFHLKYADVLFRSSLSHYEVKLN-----SSLERYOGIGPPFSCIFRIQNLGL 997
Db      926  DNRGQVTPVKYEVGFIQVSVFKEHVIIAANETIPTAINTTQIGDEVTLHYRIEKGEH 985
Qy      998  FPIHGMKMTIPIATRSNRLKLKRLDFTDEANTSCNMGNSTEYRPTVEEDLRA-- 1055
Db      986  FPMNLTQLQLYPDVTAKNTLYIIT-LSHSQNAIC-----KSSY---PVDH-LKIGSG 1035
Qy      1056  -----POLNHSNSDV-----SINCNRILVPMQEIHFHILGNLWLSLKALYKYSM 1101
Db      1036  KSYVLPKIKEPTDTIMECDTFCASINCALAPSDISQVNSL--RVWKPTIIKASHLSL 1093
Qy      1102  KIMWNAALQOFHSPRTFREEDFSROLEPESIKQ-EDWQVPMIIVGSTLGGELLALLV 1160
Db      1094  TLVVKALL-RSENSLLIRNDHQKLEPMIKISKEPPGPGSVPLMIVPLISFAGLIIALLI 1152
Qy      1161  LALRKLGFPSARRR 1176
Db      1153  FALWKAGFFRPLKKK 1168

RESULT 5
Q9QY67 PRELIMINARY; PRT; 1161 AA.
AC Q9QY67;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OC NCBI_TaxID=10116;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin.";
EL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021334; AAF21241.1; -
DR HSSP; P1215; IBBQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004985; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW INTEGRIN.
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984705E CRC64;

Query Match 19.2%; Score 1196; DB 11; Length 1161;
Best Local Similarity 29.8%; Pred No. 7e-74;
Matches 360; Conservative 205; Mismatches 514; Indels 128; Gaps 42;

Qy 7 LVVAVALSPLWPGFTDTFNMDTRKPRVIGSRITAFPGYTVQOHDISGNKVLVVGAPLETNG 66

```

Db 8 LLCWVLAASCHG----SNLDVZEP-IVFREDAASFGCTVVQF---GGSLVVGAPLEAVA 59  
 QY 67 YQKTDGVKCPVHGNCTKLNIGRVTLNVNRSERKDNMELGLSLATNPKNSFLACPLMS 126  
 Db 60 VNQTRGLXDCAPATGWCQPIVL-RSPLEAV-----NMSLGLSLVATATNAQLACGPTAQ 113  
 QY 127 HECGSSYTTGMCSSVNSFRSKTVAPALQRC-QTYMDIVIVLDGNSI--YPWVEVOH 183  
 Db 114 RACVNNVAKSGCLLIGSLQFIQAVPASPMPCEPQEMDIAFLIDSGSINQORDPAQMD 173  
 QY 184 FLINILKKFYIGGQIQGVGVQEDVHFBHNDYRSVKDVVEAASHIBORGGTETETA 243  
 Db 174 FVKALMGEF--ASTSTLSLMQYSNLKTHFTFTFEKNILDQSLVDPIVQLGL-TYTA 230  
 QY 244 FCIETARSEAF--QKGRGAKKMWLVITDGHSDP-DLEKVIQSSERDNYRVAVAVL 300  
 Db 231 TQIRTVMEFLHFKNGSRKSAKILLVITDQKRDPLEYSVPIPAADKAGIIRTAIGV 290  
 QY 301 GYNNRRGINPTEFLNEIKYIASDDPDKHFNFYTDAAALKDVIDALGDRIFSLEGT-NKNE 359  
 Db 291 DAFQE-----PTALKELNTIGSAPPQDHFVKGNPAALRSIQRLQKXIFAIEGTQSRSS 345  
 QY 360 TSFGLMSGTGSSHHVVEGVLGAVGAVDNGAVLKETSAGKVP--LRSEYLFEPFE 417  
 Db 346 SSFQHEMSQGFSSALTSDGPVLGAVGVSFSGGAP-----LYPPNTRPTFINMSQEN 398  
 QY 418 LKNHGAYLGYTVTSVVSSRQGRVYVAGAPRNFHTGKVLFTMNNRSLTIHQMERQQIG 477  
 Db 399 VDMRDSYLGYS-TAVAFWKGVHSLILGAHRHCHTKVVIPT-QEABHWRKPESEVGTQIG 456  
 QY 478 SYFGSITVDIDGQVTVLVIGAPWYNEGREKGVTVYELR--QNRFTVNGTLKQSH 535  
 Db 457 SYFGSLCVDVDKSDTDLVIGAPHYYEOTR--GGQSVFVPGVGRGWOCEATLHGEQ 515  
 QY 536 SYQNARFGSSIASVRDLNDSVNDVVVGAPLDNDHAGAIYIFHG-FRGSILKTPQORITA 594  
 Db 516 GHPWGRFVALTVLGVNDGNDLADVAIGAPGHEESRGAVIIFGASRLIEIMPSPQRVTG 575  
 QY 595 SELATGLQYFGCSIHQOLDNEDGLDGLAVGALGNATILWSPPVQINASHFBSKINI 654  
 Db 576 SGLSLRLQYFGQSLSGQDLTQGLVDLAVGAQGHVLLRLSLPLKLVLSIRFAPMEVAK 635  
 QY 655 FHRDCKRSGDATCL---NAFLCTPIFLAPHPQTTVG-IYVATMDERRTYTPRAHLE 710  
 Db 636 AVYQCWE--RTPVLEAGSATVCLTVHKGPSDLLGNVQGSVRYDLADPGRILISRAIFDE 693  
 QY 711 GGDRTFNRAVLSSGOELCERINFHVID-TADYVKEPVTFSVRYSL-----EDPDHGMPLD 764  
 Db 694 TKNCTLTGRKTLGLGDH-CETVKLLPLDCVEDAVSPILRLNPSLVDRDSASPRNLHPVLA 752  
 QY 765 DQWPTTLRVSVFPWNCNEDEHCVDPDVLDRSDLPITAMEYCORVLKPAQDCSAYTLSP 824  
 Db 753 VGSQDHITASLPPEKNCKQELCEGDL-----GISFNP 785  
 QY 825 DTTVFTIESRQRVAVAEATLENREGENAYSTVLNISQSANLOFASLI---QKEDSDSGSIRC 881  
 Db 786 SGLQVLVVGSPBLTIVTVWNEGSDSYGLTVKFFYPAGLSYRVVTGTQPHQYPLRLAC 845  
 QY 882 VNE---ERRLOKQVCMVSPFFRAKAVAPRLDSEFSKSIPL-HHLEIELAGSGSNED 937  
 Db 846 EABPAQAQEDLRSSSCSINHPIFREGAKTTFMITFDVSYKAFGLDRILLAKASSENKPD 905  
 QY 938 STKEDNVAPLRPHLKVEADVLFTRSSLSHYEVLKANSLSERYDVGIPFPFCIFRIONLGL 997  
 Db 906 TNK--TAFQLELPVKVTYVTLISGRQSDSNH-VNPSSS---HGGRRQEAHRVRYNNLSP 959  
 QY 998 FPIHGMKMITIPIATRSNRLKLRLDFTDEANTSCNIGNSTEYRTPVEE-----DL 1052  
 Db 960 LXL-AVRVNFVFPVL-----LNGVAVWDVTLSSPAQGVSCVSMKPPQNPDPFLTQI 1009  
 QY 1053 RRAPOLNHSNDSVVSINCHIRLVNQ-EINZHLIGNL---WLESLS---KALKYKSKIMV 1105  
 Db 1010 QRSVLDCSADCLHFRCDIPSLDIQDELDFILRGNLSFGVSGTQLQEKVLLVSEABITP 1069

RESULT 6  
 Q9QXH4  
 ID Q9QXH4 PRELIMINARY; PRT; 1169 AA.  
 AC Q9QXH4  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Leukocyte adhesion glycoprotein p150,95 alpha integrin subunit.  
 GN ITGAX.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Huang X., Gorski K., Tong C., Rattis F.M., Tseng S.-Y., Pardoll D.,  
 RA Tsuchiya H.;  
 RT "Isolation of Genes Selectively Expressed by Dendritic Cells."  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF211864; AAP23492.1;  
 DR HSP; P11215; 1BHQ.  
 DR MGB; MG1:96609; Itgax.  
 DR GO; GO:0008305; C:integrin complex; IEA.  
 DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.  
 DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
 DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWF\_A.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00357; Integrin\_A; 1.  
 DR Pfam; PF00092; vwa; 1.  
 DR PRINTS; PR01185; INTEGRINA.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00191; Int\_alpha; 4.  
 DR SMART; SM00327; VWA\_1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS50234; VWFA; 1.  
 KW Integrin.  
 SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;  
 Query Match 18.6%; Score 1160; DB 11; Length 1169;  
 Best Local Similarity 29.5%; Pred. No. 2.3e-71;  
 Matches 372; Conservative 215; Mismatches 498; Indels 174; Gaps 49;  
 QY 9 VAMALSLPFGTDT--FNMDTRKPRVIFGSRFAP-----FGYTVQOHDISGNKWLIVGA 60  
 Db 6 IAFLLLL--GFVSCIGFNLDAEK-----LTHFMDGAEFGHSVLQYDSS---VWVGA 53  
 QY 61 PLEITNGYOKTGDVYKCPVHGNCTKLNIGRVTLNVNRSERKDNMELGLSLATNPKNSFLA 120  
 Db 54 PKELKATNQIGLGYKCGYHTGCEPISL-----QVPPEAVNISLGLSLAATNPFWLLA 107  
 QY 121 CSPLWSHSCGSSYYTTGMCSSVNSFRSKTVAPALQRC-QTYMDIVIVLDGNSI--YP 177  
 Db 108 CGPTVHHCTRENILYTLGLCFLLSSFKQSQNPPTAQCECPKQDQDQVFLIDSGSGISST 167  
 QY 178 WYEVQHFILNLIKFFYIGGQIQGVGVQEDVHFBHNDYRSVKDVVEAASHIBORGG 237  
 Db 168 FEKMLDFKAVMSQ--LQRPSTFRSLMQFSDFYRVHFTFNFNISTSPSLSLGSRQLRG 225  
 QY 238 TETRTAFGIEFARSAP--QKGRGAKKMWLVITDGHSDP-DSPLEKVIQSSERDNYRV 294  
 Db 226 -YTYTASAKIKVITELFTTQSGARQATKVLIVIDGRKQGNLSVDSVLPWAEASIIIR 284

295 YAVAV-LGYNNRGINPETFLNEIKYIASDPDDXHFNFVTDALDKDIVDALGRIFSL 353  
 285 YAIWGVAFYN-----EHSQELKALASPSHEVFSVENFDALXDIENOLKEKIPAE 338  
 354 GT-NKNETSPGLEMSQGFSESHVVDGVLGAGVDMNGAVLKEKTSAGKVIPLRSYLK 412  
 339 GTETPSSSTFELEMSQGFSAVFTPDGVLGAGVDSFSGAFIYPS-----NRPPTFIN 393  
 413 EPPELKNHGAAYLYTVTSVSSRQGRVYVAGAFRNFHTGKVLFTMNNRSLTIHOAMR 472  
 394 MSQENEDMRDAYLGS-TALAFWKGVHSLILGAPRHQHTGKVIET-QESRHWKPEVR 451  
 473 GQQTGSYFEGSEITSDIDGVTDLVAGAPMYNEGERGKVVYELR---QHRFVYNGT 530  
 452 GTQTGSYFASGLSCVDMRDGSDTLVLGPHYHETR-GGQSVSCVMPGVGRWHCGTT 510  
 531 LKDSHYNARFSGSIASVRDLNDSYNDVVVGVAPLEDNAGAIYIFEG-FRGSILKTPK 589  
 511 LKGGQGHGKPFGRGAALTVLGVDNGDSLADVAIGAPGEENRGAVYIFEGASRODIAPSPS 570  
 590 ORITASELATGLQYFGSGIHQDLNEDGLDLAVGALGNVILKSPVVOINASHLFE 649  
 571 QRISASQIPRQYFGOSLSGGQDLTRDGLVDLAVGSGRVLRLTRPILRVSPVHTFP 630  
 650 SKINIFHRDCKRS-GRDATCLAAFLCPTPIPLAFPHFOTTV--GIRYNATMDERRYT 706  
 631 AEISRSYFEGQVAPETQLSDATVCL-HIESKTLQGLRSTVTFDLALDGRILSTRA 689  
 707 HLDGGRDFTNRVALLSGQELCERINPHV-LDTADYKVPVTSVEVSLD-----PDH 759  
 690 IFKETKTRALTRVTLGLNKH-CESVKLLPACVEDSVTITLRLNPSLVGVPISSQNL 748  
 760 GPMLLDCKPTLRVSPFPMGNEDEHCVPL-VLDARSDLPTAMEYQORVLAKPAQDCS 818  
 749 QPMLAVDDQYFTASLTFEKNCGADHIQDDLSVVFPGPDLKT-----791  
 819 AYTLSDFTTPIIESTRQVAVATELNRGENAYSTVLNISQSANLQF-----ASLIQ 871  
 792 -----LVVGSLEMLNVDVTVSGNDGDSYGTVTILFYPVGLSFRRAVAGQVFLK 840  
 872 KEDSDGS-----IECVNEERLQ---KQVNVSYPPFERAKAVAFRLDSFS-KSIF 919  
 841 KEDQWRRGSHSLHLMCDSTPDRSGLMSTSCSRHVPFRGSGQMTFLVTFDVSRAEL 900  
 920 LHLIEI LAAGSDSDNERDSTKEDNVAPLRFHLKYEADVLFRSSSLSHYEVKLNLSERY 979  
 901 GDRLLLRARVSENNVPGTEK--TTFQLELPVKY--AVVTMLSSHDPQTKYLNFTSEK 955  
 980 DGIGPPSCI---PRIONGLFPHGMKMTIPIATRSNRLKLKDLFTDEANTSCNI 1036  
 956 E---KTSVVEHFRQVNLGQDV-PVSINFVPIELK-GEAVTVMVSHPNPLTQC-- 1007  
 1037 WGNSTERYPTVEEDL---RRAPOLNHSNDVVSINCLRLVFN---OEINFHLLGNL 1088  
 1008 ---YRNLKPTQFDLLTHMQKSPVLDCSIADCLHACDI---PSLGILDELFLKGNL 1060  
 1089 ---WLRSL---KALKYKMKIMVNAALORQFHSFPFIRBEDPSRQIFEFISKEDQV-- 1140  
 1061 SFGMISQTLQKVLLESEBITFNTSVYSLQPGAEFLRAQTKTVLEM-----YKVN 1113  
 1141 PIIWIVGSTLGLLALLVLAFLKGLFERSARERREBGLD-----PTPKVLE 1188  
 1114 EVPLIVSSVGGELLLAIITALLYKAGFP---KQYKEMLEERANGQVSDGTTPQVAQ 1159

RESULT 7

Q81VA6 ID Q81VA6 PRELIMINARY; PRT; 1169 AA.

AC Q81VA6

DT 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RA Strausberg R.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC038237; AAH38237.1; .  
 DR GO; GO:0008305; C:intracellular complex; IEA.  
 DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.  
 DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWF\_A.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00357; Integrin\_A; 1.  
 DR Pfam; PF00092; vwa; 1.  
 DR PRINTS; PR01185; INTEGRIN.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS00234; VWFA; 1.  
 SQ SEQUENCE 1169 AA; 128521 MW; 178484FEFC79BB6 CRC64;  
 Query Match 18.2%; Score 1131.5; DB 4; Length 1169;  
 Best Local Similarity 28.4%; Pred. No. 2.2e-69;  
 Matches 33%; Conservative 218; Mismatches 472; Indels 201; Gaps 48;  
 QY 7 LVVAVALSMPGFTDTFNMTRKPRVPGSRTPA?-----PGYVQOHDISGNKWLUVGA 60  
 DB 10 LFTALATSL-----GFNLDTSE-----LTAFRVDSAGFGDSVVQY---ANSWVVG 53  
 QY 61 PLETNGYQKTDVYKCPVHGNCTKLNLCGRVLSNVSRKDMRGLGLSLATNPKDNFLA 120  
 DB 54 POKITAAQTGLYQCGYSTGACEPIGL-----QVPPAVNMSLGLSLATNTPSOLLA 107  
 QY 121 CSPPLMSHSGSSYTTTGMCSRVNSFRSKTVAPALQRC-QTYMDIVIVLDGNSNLYP-- 177  
 DB 108 CGPTVHEHCGRNMYLTGLCFLGPT-QLTQRLPVSRQECPRCEQDIFVLIDGSGSISR 166  
 QY 178 WVEVOHFLINILKXYIGPGQIQGVGVQYGVGVDFHNDYRSKDVVEAASHIEQ-RG 236  
 DB 167 FATMNFVRAVISQFQ--RPSTQSLMFPNFKFQTHFTFEFRFRSNPLSLASVHQLQG 224  
 QY 237 GTETRTAFIEFARSEAFQKGRGKAKVMIVITDGESH-DSPDLKVKIQQSERDNVTRY 295  
 DB 225 FVTATAIQNVHRLPHASYGARRDAAKILVITDGKEGDSLVDKVPIMADAAGIIRY 284  
 QY 296 AVAV-LGYNNRGINPETFLNEIKYIASDPDDXHFNFVTDALDKDIVDALGRIFSL 354  
 DB 285 AIGVGLAFQNRNS-----WKELNDIASKPSQEHFKVDFDALKDIQOLKEKIPAE 338  
 QY 355 T-NKNETSPGLEMSQGFSESHVVDGVLGAGVDMNGAVLKEKTSAGKVIP--LRSEYL 411  
 DB 339 TETTSSESELEMAQEGSFAVFTPDGVLGAGVSTWSGAF-----LYPPNMSPTFI 391  
 QY 412 KSPPELKNHGAAYLYTVTSVSSRQGRVYVAGAFRNFHTGKVLFTMNNRSLTIHOAM 471  
 DB 392 NMSQENVDNRDSYLGYS-TELALWKGVSQSLVLAGAPRYQHTGKAVIFT-QVSRQWRKAE 449  
 QY 472 RQQTGSYFSGSITSDIDGVTDLVAGAPMYNEGERGKVVYELRQ--NRPVYNG 529  
 DB 450 TGTQGSYFASGLSCVSDSDGSTDLVLGAPHYHETQTR-GQVSVCPPLPRGWRWNCDA 508  
 QY 530 TLKDSHYNARFSGSIASVRDLNDSYNDVVVGVAPLEDNAGAIYIFHGFTG-SILKTP 588  
 DB 509 VLYGEGQHPWGRFGAALTVLGVDNGDKLTDVVGAPGKERNRGAVYIFHGLVGPSPSPSH 568  
 QY 589 KQRIITASELATGLQYFGSGIHQDLNEDGLDLAVGALGNVILKSPVVOINASHLFE 648  
 DB 569 SQRIAGSLSRLQYFGQALSGGQDLTQGLVDLAVGARGQVLLLRPLRVLVGVSMQFI 628



```
QY 874 DSDGSECVNEER-RLQKQVNSVYPPFRAKAVAF-----RLDSEFSKSIHLHLEI 925
Db 834 RTRSS--CGRDSGATNRTCSIDLPIYRSGTTTQFLGTFRVMKNDNDSS-----NRMEI 886
QY 926 ELAAGSDSN-EROSTKEDNVAPLRFHLKYEADVLFTRSSLSHYEVKLSLSLERYDGI 984
Db 887 MITANSNNGNSMDTEVERSVYQFAV--DLAISLVAEDSVTY-----MNFSLJ--DRGP 937
QY 985 -PSCIFRIONLGLFPHGGMMKTIPIATRSNRLKLKLDLFDLDEANTS-----1033
Db 938 KPLNITYKVENSGR---KDLFVSVTLPCQTPHVLTPHTFSMHEVHHSFISSYHQIIM 994
QY 1034 CNTWGNSTERYPTVEEDLAPOLNHSNDVVSINCN-IRLVNPOBINFHLLGNLWLRS 1092
Db 995 CLLNKHLFFSP-----ELSAQVQVSTGNSLREVECSQFDLNKSSAVHFNLTADARLQN 1049
QY 1093 LKALKYK-----SKIMVNAALQRFHSPFFREDPSR-QIEFEISKQED 1137
Db 1050 VKEYESKSYFPRKUNVFSIAELNNTSLYNQTSSELKY---NPHRSQTEVKV---E 1102
QY 1138 WQVP---IWIIVGSLGGLLLALLVLAERKLGFFRSAR 1173
Db 1103 FVPPSLMLVCTGAVGGFFLLIILFLILKCGFFKRR 1141

RESULT 9
Q9J130 PRELIMINARY; PRT; 1151 AA.
AC Q9J130;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Integrin beta 2 alpha subunit.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Fathallah D.M. Sr.; Zerria K. Jr.;
RT "Cloning of the rat CD11b cDNA sequence.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; A2368593; AAR81280.1; -
DR HSSP; P1215; 1BHQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWP_A.
DR Pfam; PF01839; FG-GAP_3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWA; 1.
SQ SEQUENCE 1151 AA; 126943 MW; 8P785695D4074CA5 CRC64;

Query Match 17.6%; Score 1096.5; DB 11; Length 1151;
Best Local Similarity 28.8%; Pred. No. 5.8e-67;
Matches 357; Conservative 203; Mismatches 494; Indels 185; Gaps 49;

QY 7 LVNAWALSIPGFTDTFMOTRPRVPGSRTAFPGYVQOHDISGNKWLAVGAPLETNG 66
Db 5 VLLATVLTLCG-----FNLDTEPMTPQENASG-FGOSVIQ---LGTRVVVAAPQEVKA 56
QY 67 YQKTDGVYKCPVHICNTCKNLGRVLSNYSERKDNRLGLSLATNPKDNSFLACSP LWS 126
Db 57 VNQTGALYQCDYSNRCDDIPL-----QVPEAVNMSGLSLAAATVTPVQLACGPTVH 110
```

[illegible]



|   |   |  |
|---|---|--|
| QX  | NCBI_TaxID=10090;   |  |
| RN  | [1]   |  |
| RP  | SEQUENCE FROM N.A.  |  |
| RC  | STRAIN=C57BL/6J; TISSUE=Spleen;                                       |  |
| RA  | Ma R.2.; Teuscher C.;   |  |
| RL  | Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.               |  |
| DR  | EMBL; AF065901; AAD25884.1; -   |  |
| DR  | HSP; P20701; ILPA.  |  |
| DR  | MED; MGI:96606; Itgal.  |  |
| DR  | GO; GO:0008305; C:integrin complex; IEA.                              |  |
| DR  | GO; GO:0004895; P:cell adhesion receptor activity; IEA.               |  |
| DR  | GO; GO:0007160; P:cell-matrix adhesion; IEA.                          |  |
| DR  | InterPro; IPR000413; Integrin_alpha.                                  |  |
| DR  | InterPro; IPR002035; VWF_A.   |  |
| DR  | Pfam; PF01839; FG-GAP; 3  |  |
| DR  | Pfam; PF00357; integrin_A; 1.   |  |
| DR  | Pfam; PF00092; vwa; 1   |  |
| DR  | PRINTS; PRO1185; INTEGRINA.   |  |
| DR  | PRINTS; PRO0453; VWFADOMAIN.  |  |
| DR  | SMART; SMC0191; Int_alpha; 5.   |  |
| DR  | SMART; SMC0327; VWA; 1.   |  |
| DR  | PROSITE; PS00242; INTEGRIN_ALPHA; 1.                                  |  |
| DR  | PROSITE; PS02034; VWF_A; 1  |  |
| SQ  | SEQUENCE 1160 AA; 128127 MW; A33C531B139F1FAD CRC64;                  |  |
| Query Match   |   |  |
| Best Local Similarity 28.1%; Pred. No. 1.3e-64;                     |   |  |
| Matches 358; Conservative 204; Mismatches 451; Indels 262; Gaps 55; |   |  |
| QY  | 4 PRGLVVAWAL--SLWPGFTDTENMDTEKPRVIPSRTAFPGYTVQQHDISGNKWLVVGAP 61      |  |
| DB  | 8 PRLLLLGLQFAKAW-----SYLDTPTQSFQAQGRHFGYQLQIE-----DGVVVGAP 58         |  |
| QY  | 62 LETNGYKQTDGVYKCPVIHGNCTKLNIGRVTLSNVSRKONMRLGSLATNPKNDFLAC 121      |  |
| DB  | 59 GEGD---NTGGLVHCRTSSPQQPVSLLH-----GSNHTSKYLGMTLATDAAGSLLAC 108      |  |
| QY  | 122 SPLMSHRCGSYVYTCMSRVNSNFRSKTV-----APALQRC-QTYMDIVIVLDGNS 173       |  |
| DB  | 109 DPLGSRDCDQNTYLSGLC-----YLPQSLGPMQNRPAYPECMKGVDLVFLDGSQ 162        |  |
| QY  | 174 SI--YPMVEVQHPLINKFYIGPQOIQGVVQYGEDVVEHFLNDYRSVK-----DVV 226       |  |
| DB  | 163 SLDRKDFEKILEFMKDWK--LSNTSYQFAAQVQFDCRTEFTFLDY--VKQKQNPDLV 218     |  |
| QY  | 227 EAASHIQRGCTE-----TTAFGIEFARSEAF--QKGRKGAKKMIIVITDGSHPD 279        |  |
| DB  | 219 -----LGSVQPMFLITMFRAINVYVAHVFKESGARPDATKVLIVITDGERSDGN 270        |  |
| QY  | 280 LEKVIQSERDNTRYAVAVLGYNNRGINPETHINEIKYIASDPDDKHFFNVTDI-AAL 338     |  |
| DB  | 271 I-----SAADIIIRYIIGIKHP--VSVQKQKTLH---IFASEFVEE-FVKILDTFEKL 318    |  |
| QY  | 339 KDIVDALGDRIFSEGTNKEB-TSPGLEMSQTGFSSHVVEDGVLGAVGYDNGAVLKE 397      |  |
| DB  | 319 KDLFTDLQRIYAIETGNRQDLTSFNMLSSSGISADLSKGHAVGVAGAKDW----- 371       |  |
| QY  | 398 TSAGKVTPLRE-----SYLKEFPEELKNHGYLGTYVTYSVVSRRQGVYVAGAPRNFHTG 452   |  |
| DB  | 372 --AGFLDLREDLQAGTFVCGEPILTSVRCGLYGTGYVAMTSSRRPILLAGAPRYQHV 429     |  |
| QY  | 453 KVILFTMN-----NRSITIHQAMRGQOIGSYFGSEITSDVIDGQVTDVLLVAGPYFN 507     |  |
| DB  | 430 QVLLFQAPAGGRVNTQ-----QKIEGTQIGSYFGELCSVDLDQDGAELLIGAPLFFG 485     |  |
| QY  | 508 EGRERGKVVYVELPQNFVYNGTLKDSHTYQNARFGSSIASVRDLNDSYNDVVVVGAPLE 567   |  |
| DB  | 486 EQR--GGRVFTYQRRQSLEFMVSELOQDPGPGPLGRFGAAITALTDINDRLTDVAVGAPLE 544 |  |
| QY  | 568 DNHAGAIYFHGFRGSIKTPKORITASELATGLQYFGCSIHGOLDNEDGLDLAVGAL 627      |  |
| DB  | 545 EQ--GAVYIFNPKGGLSPQPSQRIQGAQVFPFGIRWFGRSIHGVKDLGGRLANVAVGPE 602   |  |
| QY  | 628 GNAVILWSRPVVVINASLHFEPSKINITHRCKRGRDATCLAAPL--CFTPIFLAPHFQ 685    |  |
| DB  | 603 GRVVVLSRPVVVDVTELSFSEEEIPVHEVECSYARBQKHGKVLKACFRIKPLTPQFQ 662     |  |
| QY  | 686 -TTVGIRYNATMDERRYTTPRAHLDEGGDRFTNRRAVLLSSGOELCERINHY-LDTADTV 743  |  |
| DB  | 663 GRLLANLSYTLQLDGHMRSEGLPFGSHLSGNTSI--TPDKSCLDFHFPFICIQDLI 720      |  |
| QY  | 744 KVPFSEVSELDPPDHPMLDDG-----WPT--TLRVSVPFMNGCNDEHCVDPDLVD 794       |  |
| DB  | 721 SPINVSINFSLLEBSGTPRQKGRAMQPIILRSIHVTKEIIFPKNCKEDKCEANLTL 780      |  |
| QY  | 795 --ARSDLPAMBYCORVLARKPQDCSAYTSLSDTTVTTFIESTRQRVAVPATLENGENAY 852   |  |
| DB  | 781 SPARGG-----PLRLMSSASL-----AWEWTLNSGSDAY 810                       |  |
| QY  | 853 STVLNLSQSANLQF--ASLIQ-----KEDSGSIECVNEERRLOQV--CNVSYPPF 900       |  |
| DB  | 811 WYRLDLDPRLGSLFRKVMQLPHSRPVSCBELTEGS-----SLLTYTLKCNVSSPIF 863      |  |
| QY  | 901 RAKAKVAFR-----LDSEPSKSIPLH--HLEIELAAGSDSNERDSTKEDNVAPLREHL 951    |  |
| DB  | 864 KAGQEVSLQVMFNTLLNSGSDFWELNGTVHCE-----NENSSLQEDNSAATHIPV 914       |  |
| QY  | 952 KYEADVL-----FT-----RSSSLSH-YEVKLNSS-----LERYDGIQPPF 986           |  |
| DB  | 915 LIPVMLTKBQENSTLYISFTPKGPKTQQVQVHYQRIQPSAYDHNMPTLEALVGVPRH 974     |  |
| QY  | 987 SCIFRITQNLGLFPFHGMKMITIPIATRSNRLKLRLDPLTDEANTSCNINGSYETRT 1046    |  |
| DB  | 975 S-----EDLIT-----YTWSVQTDPLVT 992                                  |  |
| QY  | 1047 PVEDILRAPQLNHSNDSVSNICNIRLVNQNFINPHLLGNLWL-RSLKALKYKSM--KI 1103  |  |
| DB  | 993 CHSEDLKRPSSBAEPLCGVQPRCPF--VFRWEILIQVTGVTELSKBIKASSTLSCSSL 1050   |  |
| QY  | 1104 MYNAALQROFHSPFIFREDPSRQIEFEISKQEDQVPIIIVGSLGGLLLALLALVAL 1163    |  |
| DB  | 1051 SVSENSSKHFH--LYGSKASEAQVLVKVLDLHEKEM-LEVVVLSGIGGLVLLFLIFLAL 1106 |  |
| QY  | 1164 RKLGHF-RSARRRE 1177  |  |
| DB  | 1107 YKVGFFKRYLKEKE 1121  |  |
| RESULT 12   |   |  |
| Q98TF0  | PRELIMINARY; PRT; 1187 AA.  |  |
| AC  | Q98TF0;   |  |
| DT  | 01-JUN-2001 (TrEMBLrel. 17, Created)                                  |  |
| DT  | 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)                     |  |
| DT  | 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)                   |  |
| DE  | CD11-2.   |  |
| GN  | CIA2;   |  |
| OS  | Cyprinus carpio (Common carp).  |  |
| OC  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;     |  |
| OC  | Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  |  |
| OC  | Cyprinidae; Cyprinus.   |  |
| OX  | NCBI_TaxID=7962;  |  |
| RN  | [1]   |  |
| RP  | SEQUENCE FROM N.A.  |  |
| RC  | TISSUE=Peritoneal exudate cells;                                      |  |
| RA  | Kimura M., Fujiki K., Nakao M.;                                       |  |
| RT  | "Molecular cloning of a leukocyte integrin from the common carp.";    |  |
| RL  | Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.               |  |
| DR  | EMBL; AB048537; BAB39135.1; -   |  |
| DR  | HSP; P20701; ILPA.  |  |
| DR  | GO; GO:0008305; C:integrin complex; IEA.                              |  |
| DR  | GO; GO:0004190; P:aspartic-type endopeptidase activity; IEA.          |  |
| DR  | GO; GO:0004895; P:cell adhesion receptor activity; IEA.               |  |
| DR  | GO; GO:0007160; P:cell-matrix adhesion; IEA.                          |  |
| DR  | GO; GO:0006508; P:proteolysis and peptidolysis; IEA.                  |  |
| DR  | InterPro; IPR001969; Asprotease_AS.                                   |  |

|   |  |  |
|---|--|--|
| DR  | InterPro; IPR000413; Integrin_alpha.                                 |  |
| DR  | InterPro; IPR002035; VWF_A.  |  |
| DR  | Pfam; PF01339; FG-GAP; 3.  |  |
| DR  | Pfam; PF00357; Integrin_A; 1.  |  |
| DR  | Pfam; PF00092; vwa; 1.   |  |
| DR  | PRINTS; PR01185; INTEGRINA.  |  |
| DR  | PRINTS; PR00453; VWFADOMAIN.   |  |
| DR  | SMART; SM00191; Int alpha; 5.  |  |
| DR  | SMART; SM00327; VWA; 1.  |  |
| DR  | PROSITE; PS00141; ASP PROTEASE; 1.                                   |  |
| DR  | PROSITE; PS02334; VWF_A; 1.  |  |
| SQ  | SEQUENCE 1187 AA; 131778 MW; 85EDC7CA8B6B1C59 CRC64;                 |  |
| Query Match   |  |  |
| Best Local Similarity 26.0%; Score 1045; DB 13; Length 1187;        |  |  |
| Matches 323; Conservative 245; Mismatches 463; Indels 210; Gaps 50; |  |  |
| Qy  | 13 LSLMPGFTTDMIDTKPRVIFGSRRTAFPGYVQVQHDISGNKMLVVGAPLENGYQKTKD 72     |  |
| Db  | 22 LFMASLSAFNIDIEHLRFNGTPDFGYSYQTEFGNRKQIIVGAPLEGN---SAGE 78         |  |
| Qy  | 73 VYKCPVIHGNCTKLNLRVTLNVSEKKNMR-LGLSLATNPKNDSFLACSLWHSBGCS 131      |  |
| Db  | 79 MYSCATLQSCARLQ-----RPGSESVRFPGMSAAVS--SAALTSCSYFAHECDG 127        |  |
| Qy  | 132 SYTTGMCGRVSNFRFSKVPALQRC-QTYMDIVIVLDGNSIYPMWEV---QHPLIN 187      |  |
| Db  | 128 NSYLANGVCYQFNSSLQAVSNFTAAAYQECSEKREVNLFDFGSSMKA-VEFDMKNFKIKD 186 |  |
| Qy  | 188 ILKKFVIGQIQGVGVQVQGVQGVQGVQGVQGVQGVQGVQGVQGVQGVQGVQGVQGV 245     |  |
| Db  | 187 VMKK--LSNNSIKFAAQVQSTEITVDFNDYQNGSAEKKMKRMK----SLNTYKA 240       |  |
| Qy  | 246 IFARS--EAFQGRKGKAKVMIVITDGEHSDSPLEKYQQSERDNTVRYAVAVLGY 302       |  |
| Db  | 241 INVVLKNVLSVSSGADPAQAKALVIIIDGDFSDNDY-NILNICDEQNILAYIIGV--- 296   |  |
| Qy  | 303 YNRGINPTFLNEIKYIASDDPKHFFNVYDEALKDVIDALGRIPSLSEGTNK-NEFS 361     |  |
| Db  | 297 -----GKVDITLTQLAAEPKLANTFYIQYNGLKLGLDLQKVIYNIEGSKAAGRD 349       |  |
| Qy  | 362 FGLEISQTFSSHVEDGVLGAVGAYDMNGAVLKETSAKVPIRESYKLFEPPELKNH 421      |  |
| Db  | 350 RQKELSQSGFVYQBEVIVSVGSNDWRGALYVWMSG--KFRQT---SITPAVYNK 404       |  |
| Qy  | 422 GAYLGYTVTSVSSRQG-RVYVAGAPRNHTGKVLFTMHNRSITHOAMRGOQIGSYF 480      |  |
| Db  | 405 DSYMGY--STVLGMRHGVSLFSGAPRAEHTGLVTLFT-KNESITWTWENINGEIQISYF 461  |  |
| Qy  | 481 GSEITSVDIDGQVTDVLVGPMTY-NEGRERGKVVYVELQNRFPVYNGTLKDSHSYQ- 538    |  |
| Db  | 462 GASLSDLDVSDGSDFLVGLAPLFYQSPRAEGRLVY-----TLSSQISQKT 510           |  |
| Qy  | 539 ----NARFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGRGSIKTPQRITA 594     |  |
| Db  | 511 LQSTTGRFATSLASLKLDNGDGLSDVAVGAPLENGSVVYILGDTGTHINPEHAQPIPA 570   |  |
| Qy  | 595 SELATCLOYFGCSIHGOLDNEDGLDLAVGALGNVILWSPPVQVNASLHPEPSKINI 654     |  |
| Db  | 571 RSVLPGLQFGVSLSGQMDNDNDLPDIVIGQGVILLNARPVSVSAQLSFNPMELSL 630      |  |
| Qy  | 655 FRDCKRGRDTCATLAAFLCF--TPIFLAPHQITTT-----VGIRYNATMDERTYTPRA 706   |  |
| Db  | 631 NYFECGGSN-----AFNAPNLTSCTVTERTSSGSLKLVNSLNLNDVVGMRG 683          |  |
| Qy  | 707 HLD--EGGDRFTNRAVLLSSGQELCERINPHVLD-TADYVKPVTFSVEYSLDDPHG--- 760  |  |
| Db  | 684 FPDPMDSRSRTQQSVLLDSSGS-CSNFSIFMLRCVADTVSPFKIRMNFSQTMLSGNLS 742   |  |
| Qy  | 763 PMLDGCWPTTLRVSPVFMNGCNEDEHCVEDLVLDARSDLPTAXEYQCVLRKPAQDCSAY 820  |  |
| Db  | 743 AVLDIQSRTEBYEVELFQRCNCSNS-CVADLKLN----- 775                      |  |

|  |   |                           |
|--|---|---------------------------|
| Qy   | 821 TLESDTTVFIIESTRQVRVAEATLENRGENAYSTVNISQSANLOFASLIQKEDSGSIB 880      |                           |
| Db   | 776 -PSTNDTLVVEN-QAHTVVLVSLANPGDDSYNTSIVLHYPEGLSLKFDAIKPRTKSS 833       |                           |
| Qy   | 881 CWNEER-RLQOVNCNVSYFFRAKAKVAF-----RLDSEFSKSIPLHLHLSIELAAGSD 932      |                           |
| Db   | 834 CGDRDSGATNRTTCSINLPVYKSGTTTQFLGTFRTVKWDYDWS-----DRMEMTITANS 888     |                           |
| Qy   | 933 SNER--DSTKEDNVAPLRPHLYKAEADVLPTRSSLSHYEVKLSLERYDGIQIP-PPSCI 989     |                           |
| Db   | 889 NNGNMSDMSVRRSI-PVQFAV--ELAISLVAEDSVTY-----LNFSL-----DRGPKPLNII 938  |                           |
| Qy   | 990 FRIQNLGLPPIHGMKTIPIATRSNGRLKJROFLTDEANTSCNMGNSTEYRPTPVE 1049        |                           |
| Db   | 939 YKVNVLGL--KGLPVSVTLSPCQTHVTLTPHNFQMGEVHS----- 979                   |                           |
| Qy   | 1050 EDLRAPOLNHSNDSVSNIRLVPNOE---INFHLG-LNWL-----RSLKALKYKSM 1101       |                           |
| Db   | 980 -----FISSYHQIIMCLLNKHLFPSPELSAVQTRTTGRSLMWCVSSISTGEIPRSSV 1031      |                           |
| Qy   | 1102 KIMVNAALQ--RQFHSPPIFRE-----EDPSROIETPEISK-----Q 1135               |                           |
| Db   | 1032 NIMAEAVLQNVKEYESKYSFYEFPRDEHVFNTSABELFNFTSYNQSTGLKYNPHRSQTEVK 1091 |                           |
| Qy   | 1136 EDMQVP-----IMIVGSTLGGULLLALLVIALVKLGFFRSAR 1173                    |                           |
| Db   | 1092 VEFVIPPRLIVCTGAVGGFFPLIILILLKCGFFKRR 1132                          |                           |
| RESULT 13  |   |                           |
| ID   | Q06271  | PRELIMINARY; PRT; 780 AA. |
| AC   | Q06271;   |                           |
| DC   | 01-NOV-1996 (TREMELrel. 01, Created)                                    |                           |
| DT   | 01-NOV-1996 (TREMELrel. 01, Last sequence update)                       |                           |
| DT   | 01-JUN-2003 (TREMELrel. 24, Last annotation update)                     |                           |
| DE   | Integrin alpha-2 subunit (Fragment).                                    |                           |
| OS   | Xenopus laevis (African clawed frog).                                   |                           |
| OC   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;       |                           |
| OC   | Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;            |                           |
| OC   | Xenopodinae; Xenopus.   |                           |
| OX   | NCBI_TaxID=8355;  |                           |
| RN   | [1]   |                           |
| RP   | SEQUENCE FROM N.A.  |                           |
| RA   | Meng F., Desimone D.W.;   |                           |
| RL   | Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.                 |                           |
| RS   | [2]   |                           |
| RP   | SEQUENCE OF 95-168 FROM N.A.  |                           |
| RX   | MEDLINE=94008528; PubMed=8404528;                                       |                           |
| RA   | Whittaker C.A., Desimone D.W.;  |                           |
| RT   | "Integrin alpha subunit mRNAs are differentially expressed in early     |                           |
| RT   | Xenopus embryos.;"  |                           |
| RL   | Development 117:1239-1249(1993).  |                           |
| DR   | EMBL; L43058; AAA69770.1; -   |                           |
| DR   | EMBL; L10186; AAA6246.1; -  |                           |
| DR   | PIR; I51524; I51524.  |                           |
| DR   | GO; GO:0008305; C:integrin complex; IEA.                                |                           |
| DR   | GO; GO:0044895; F:cell adhesion receptor activity; IEA.                 |                           |
| DR   | GO; GO:0007160; P:cell-matrix adhesion; IEA.                            |                           |
| DR   | InterPro; IPR000413; Integrin_alpha.                                    |                           |
| DR   | Pfam; PF01839; FG-GAP; 3.   |                           |
| DR   | Pfam; PF00357; Integrin_A; 1.   |                           |
| DR   | PRINTS; PR01185; INTEGRINA.   |                           |
| DR   | SMART; SM00191; Int alpha; 4.   |                           |
| DR   | PROSITE; PS00242; INTEGRIN_ALPHA; 1.                                    |                           |
| FT   | NON_TER 1   |                           |
| SQ   | SEQUENCE 780 AA; 87017 MW; 9518B18C2B6F637 CRC64;                       |                           |
| Query Match  |   |                           |
| Best Local Similarity 16.7%; Score 1040; DB 13; Length 780;        |   |                           |
| Matches 257; Conservative 174; Mismatches 297; Indels 90; Gaps 22; |   |                           |
| Qy   | 385 VGAYDWNGAVLKEKTSAG-KVIPLESYLKEFPPEL--KNHGAYLGYTVTSVSSRQGRVY 441     |                           |



Db 1 VGYDWSGTVVHESKQFSFP-----YKAFKVLHNRNQSYYLGSV-AVINKNKSVHF 54  
Qy 442 VAGAPRHNHCKVILFTMHNRSLTHQAMRGQOIGSYFGSEITSDVIDGCGVTVLLVG 501  
Db 55 VAGAPRTDYTGQVWYNITSQGVNITQSGVNIQTQGEQIGSYFGSVLCSDVNRDSITDVLWG 114  
Qy 502 APWYFNE-GERGKVVYVLRQNRFPVYNGTLKDSHSYQNAFGSSIASVRDLNODSYNDV 560  
Db 115 APTFMNEYKEEGVYMFSDGILVQREOLEGPKSLENTFGSAIVELSDIDLQYNDV 174  
Qy 561 VVGAPELBNHAGAIYIPHGFRGILKTPKORIITASELATGLQYFGGSIHQGLDLNBDGLI 620  
Db 175 IVGAPELBNHAGAIYIPHGFRGILKTPKORIITASELATGLQYFGGSIHQGLDLNBDGLI 620  
Qy 621 DLAVGALGNVILWSPVVOINASLHPEPSKINIFHRDCKSRGDATCLAAFCPTPIEL 680  
Db 235 DVSVAIGVQVILWSPVVOINASLHPEPSKINIFHRDCKSRGDATCLAAFCPTPIEL 680  
Qy 681 APHFTOTTGIRINATWD-----RRYTPRAHLEGGDRFTNRNRAVLSSGOELCERINFV 736  
Db 286 RPANSQNSVAVTNTATLADALLSRVTSRQGFRESNDRFLQRNIVIGTPTSCFEHV-FNV 344  
Qy 737 LOTADYVXPTVPSEVSELEDDHGMPLDDGMPITLRSVPFVWNGCKNEDEHCVPLDLAR 796  
Db 345 QETSDSENALRINVAQPGSSPVLPYSSSTENFIFPLKDCGDKICISDLSLQIQ 404  
Qy 797 SLDPTAMEYCORVLRPAQDCSAYTLSPDTTPIIESTQRVAATELRNENAYSTVL 856  
Db 405 -QIPSDIK-----SPVIVSKDRLLFQVTLNKLNAVNTL 441  
Qy 857 NISQSANLOPAS-----LLOKEDSGSTECVNEERLOKQVNVSPFPRAKAV 906  
Db 442 NAIPELNFASSTIPRDTGTEVLCQNTQNSVCL-----IGFPIKDSQKV 489  
Qy 907 APRLSEFSSKIFLHLEIBLAAGSDSNRDSYKEDNVAPLRPHLYEADVLTFRSSLS 966  
Db 490 IPDIWFENLNLONTAFIVFQATSES-REAREEDNSVITLPVQYDAEHLTRVTIN 547  
Qy 967 HYEV-----KLNSLERYDGIQPPSCIFRIONGLPFIHGMKMTIPIATSGNRLKL 1022  
Db 548 FYEVFPLGNIPSVVNTDDIGLEFNFTKV-STGYPLKTAVALISINVOFTKSNPLLYI 606  
Qy 1023 RDLTDE-ANTSCNINWNSTE-----YRPTPVEEDLRAPQLNHSNDVVSINC---NIR 1073  
Db 607 TSVQVQMTGVSSNRVINPAHVEKIPYKPSFKEENLKAHVNLNCSARCAIKCFLENLQ 666  
Qy 1074 LVPNQENFHLGNLWLSLKALKYKMKIMVNAALQROFHPSPFIFREBDDPGRQIEFTIS 1133  
Db 667 LREBYTVN--VSTRINWGTFAAAHFTQIELVANAKLETIDSEIFVVSNDV--MPLTIT 722  
Qy 1134 KQ-EDQVPTIIVGSLGGLLALLLVLALRLKGFPR 1170  
Db 723 KPGKSEVPICVIGSAIAGLVAALLVALLWKLGFHK 760

## RESULT 14

Q96HB1 PRELIMINARY; PRT; 1086 AA.  
ID Q96HB1;  
AC Q96HB1;  
DT 01-DEC-2001 (T-EMBLrel. 19, Created)  
DT 01-DEC-2001 (T-EMBLrel. 19, last sequence update)  
DT 01-JUN-2003 (T-EMBLrel. 24, last annotation update)  
DE Hypothetical protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymph;  
RA Strausberg R.;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC008777; AAH08777.1; --  
DR CO; GO:0008305; C:inhibitor complex; IEA.  
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.  
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
DR InterPro; IPR000413; InterPro; Integrin\_alpha.  
DR InterPro; IPR002035; VFW A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00992; vwa; 1.  
DR PRINTS; PRO1185; INTEGRINA.  
DR PRINTS; PRO1453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS0234; VWA; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 1086 AA; 119223 MW; F6PF2546E8C632F9 CRC64;  
  
Query Match 16.4%; Score 1018.5; DB 4; Length 1086;  
Best Local Similarity 27.8%; Pred. No. 1.4e-61;  
Matches 350; Conservative 184; Mismatches 420; Indels 303; Gaps 53;  
  
Qy 8 VVAMALSLWPGF-----TDTNMDTRKPRVIPGSRTA-PFGYTVQOHDISGNKMLVVGAP 61  
Db 6 ITVAMALLSGPFPFAPASSYNLDVGRGSRSPPRAGRHFGRVQLQ---VGN-GVIVGAP 61  
Qy 62 LETNGYQKTDVYKCPVIHGMCTKLNLRVTL--SNVSRKDNMRLGLSLATNPKDNQSL 119  
Db 62 GEGN---STGSLYQCQSGTGHCLP-----VTLRGSNYTSK---YLGMTLATDPTDGSIL 109  
Qy 120 ACSPLWSEHGSSYYTTCMCSEVNSNFRFSKTVAPALQRCQTYMDIVIVLDSGNSIYPMV 179  
Db 110 FNA-----  
Qy 180 EVQHFLINILKFFYIGPQIQVGVVYQGEDVVEHPLNDYRSVKDVAASHIEQRGGTE 239  
Db 113 -----VQPTSYKTEPDSYVYKRPDALLKHVEM-LLL 147  
Qy 240 TETAFGIFBARSEAFQK--GGRKGAKYMIITGESHDSPLDKVIOQSRDNYTVAY 297  
Db 148 TMTGAINVATEVREELGARPDATKVLIIITDEATDSGNDAA-----KDIRIIV 201  
Qy 298 AVLGYNRRGINPETFLNEIKYIASDPDRHFFNVTD--AALKDVIDALGRIFSLGNT 356  
Db 202 GI-GKHFTQESQET---LHKFASKPASE-FVKILDTFELKDLFTELQKIVYIEGTS 255  
Qy 357 KYE-TSFGLEMSQTFSSHVEDGVLLGAVGAYDNGAVLKETSAGKVIPLR-----ESY 410  
Db 256 KODLTSFNMELSSGISADLSRGHVAVGAAGKDW-----AGGFLDLKADLQDDTF 306  
Qy 411 LKEFPEELKHGAVLYGTVTVTSVSRQGRVYVAGAPRNFHTGKVLFTM-----HNNRSL 465  
Db 307 IGNEPLTEVRAGLYGTVTVLPSRQKTSLLASGAPRYQHMGVRLLPQEPQGGHWSQV 366  
Qy 466 TIHQMRQQQIGSPFGSEITSDIDGCVTVLVLGAPMYFNGRERKQVTVYELQRNRP 525  
Db 367 TIH-----GTQIGSYFGGELCGVDQDQGETELLIGAPLFYGBQR-GGRVFIYORRQIGF 421  
Qy 526 VYNGTLKDSHSYQNAFGSSIASVRDLNODSYNDVWVGAPELBNHAGAIYIPHGFRGIL 585  
Db 422 EVSELOQDQGPYPIGRFGEAITALTIDNGDLVDVAVGAPLEEQ--GAVYIFNGRGGSL 479  
Qy 586 KIPKQRIITASELATGLQYFGGSIHQGLDLNBDGLIDLAVALGNVILWSPVVOINASL 645  
Db 480 PQPSQRIEQTQVLSGIQWFGRSIEGVKXOLEGDLADVAVGAESQMVILSSRPVVDVMTLM 539  
Qy 646 HFESKINIPIREDCK-----RSGRDATCLAAFCPTPIELAPHFTOTTV-GIRYNATM 697  
Db 540 SPSFAEIPVHEVCSYSTSNOKGWNIT-----ICFOIKSLIIFQFQGLVANITYTLQL 594  
Qy 698 DERRYTPRAHLEGGDRFTNRNRAVLSSGOELCERINFHV-LDTADYVXKPTVPSVEYSLD 756  
Db 595 DGHR-TRERGLFPGGRHELNRNIATVTSMS-CTDFSHPFVQVDLISPINVLSNFSLWE 652

```

QY 757 PDHGP---MLDDGMPFTLAVS-----VFENNGCNEDEHCVPLVLDARSDLPTAMEYC 806
Db 653 EGTFRDQAGKD:PIILPSLHSETWEIPFEKNCEDKKCEANL----- 697
QY 807 QVILKAPQDCSAYTLSPDTTTFIESTQRVAVENATLENGENAYSTVLTNSQSANTQF 866
Db 698 -RVSPSPARSARLRLTAFAS-----LSVELSLNLEEDAYVWQLDLHPFPGLSF 745
QY 867 ASL-IOKEDSDGSIEC--VNEERRLQKOV--CNVSYPPFRARAKVAFLI-----DSBFS 915
Db 746 RYVEMLKHSQIPVSCSEELPEBSRLLSRALSNCVSPIPFAGHSVALQNMFTLVNNSWG 805
QY 916 KSIFLHLLEIELAAGSDNERSTKEDNVAPLRFHLKYEADVL-----FTRSS 963
Db 806 DSVELH-----ANVTCCNEDSDLEDNSATTIIPILYPINILIOQEDSTLYVFTPKG 859
QY 964 SLSH-----YEVKLANS-----LERYDGIQPPFSCIFRQNLGLPIHGMKMLIIP 1010
Db 860 PKHQVKKHYQVRIQPSIHDDHNIPTLEAVVGPQP-----PSEG-----P 899
1011 IATRSNRLLLKLRDLFTDBANTSCNIWGNSTYRPTPVE-EDLRAPQLNHSNDVWSIN 1069
900 ITHQ-----WSVQMB-PPVPCHYEDLERLPDAAEPCFALER 936
QY 1070 CNIRLVNQEINHLGNLWL-----RLKALKYKMKIMYNALQRFHSPFIFRBE 1123
Db 937 CPV--VFROEILVQVIGTLELGEIEASSMFL-CSSLISISFNSS--KHFI---LYGSNA 988
QY 1124 PSRQIEFEISKEDQWPTIWIIVGSLGILLALLLALRLXGLFF-RSARRRRPPG 1179
Db 989 SLAQVVMKVDVYVYKQK-LYLVLSGIGSGLLLILLIFVLVYKGFKNLKEKMEAG 1044

RESULT 15
O88341 PRELIMINARY; PRT; 1167 AA.
ID O88341
AC O88341
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DE Integrin alpha E2 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020046; AAC23663.1; -.
DR HSSP; P11215; 1BHQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADONAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PSS0234; VWFA; 1.
FT NON TER.
SQ SEQUENCE 1167 AA; 128593 MW; E3FED7E18B5CBBEC CRC64;

```

```

Query Match 15.7%; Score 979; DB 11; Length 1167;
Best Local Similarity 26.3%; Pred. No. 9e-59;
Matches 337; Conservative 212; Mismatches 479; Indels 252; Gaps 48;

QY 10 AWALSMPGFTDTTNDTRKPRVPGSKTAFEGTVVQOHDISGNKWLUVGAPLETNGYQK 69
Db 25 SWVTALQPGAPSVLN-----SLHQDPSNQTCILLVARRSN--RN 63
QY 70 TGVVYKCVIHGNCNCKL--NLGRVTLNSVSEKDNMRGLSLATNPKNSFLACSPWS 126
Db 64 SAVIYQC-----ATSIIPDEIGQCPVEHILMPKRYQ-GVTLVRN--HNGVLVCIQVS 114
QY 127 HECGS-SYTTTGMCSRNSNF----- 146
Db 115 RKPSRLNSELTCACSLTPNLDLQAQVFSDEGLVLDLGASVNSGDYSGSGSTGEETK 174
QY 147 ---RSKTVAPALQRCQTYMDIVIVLGSNSIYP--WVEVQHFLINILKVIYIGQIQV 201
Db 175 SARLRQAVBEBEERAGTBIAIVLDSGSIERSDFQAKDFISTMWNRYEKFCECNF 234
QY 202 GWVOYGEDVVEHFLNDYRSVKVVEAASHIEQRGGTETRTAFGIEFARSAP--QKGR 259
Db 235 ALVQYGGVIQTEFLLDSRDINASLAKVQSVQVKEV-TKTASAMQHVLNDIFIPSRGR 293
QY 260 KGAKYMWIVITDGHSHSP-DLEKVIQOSERDNTTRAVAVLGYNRRGINPETFLNRIK 318
Db 294 KALKKMWVLTGDI FRDPLNLATTVISSKQGVVRAFGVGNAPF-----NNNTY-RELK 348
QY 319 YIASPDQDKHFNVYDDEAALXDIVDALCDRIPLSEGTNKNKTSFGLMSQTFGSHVVED 378
Db 349 LIASDPKAAHTFKVTNYNSALDGLSLKQRIIHEGTVGDTLQY--QLAQTFSAQILDK 406
QY 379 G-VILGAVGAYDN--GAVLKETSAGKVIPLRESYKLPPEBELKNHGYLGTVTTSVSSR 436
Db 407 CQVLLGTGAFNWSGGALLYNTQNGRGRFLNQAKEDFRAA---QYSYLGYSV-AALHKA 462
QY 437 QGRVYVAGAPRNFHTGKVIPLFTMNNRSLTTHQAMRGQOIGSYFGSETTSVIDGDGVD 496
Db 463 HGVSYVAGAPRKLKRGAVFELQKEDGEB-TFMKRIEBSQMSYSGSVLCPVDINWDGTD 521
QY 497 VLLVGA PMYFNNEGRGKVVYVELRQN----RFVYNGTLKDSHSHYQNAFSGSIASVRL 552
Db 522 PLLVAAPFYHIRG--EGRGVYVYRVEHQDAPPSLVY--TLSGYPGLTSSRFQFAMAAGDI 578
QY 553 NODSYNDVNVGAPLEDNEA-----GAIYFPHGFRGSLTKTPKORITASELATGQYFGC 606
Db 579 NQDKFTDVAIGAPLEGFAGDGASYSVYINGHSGGLHASPSQOIRASSVALGLYFPM 638
QY 607 SIHQQLDNLBGLIDLAVGALGNVILWSRPVQINASHLHFEPSKINIPIHRDCKRSGDA 666
Db 639 SVSGGLDFSGDDLADITVGSQDVAVLRSPVVDLTVSMFTPDALPMAFKD----- 690
QY 667 TCLAAFLCFTPIFLAPHFQTTVCIIRYNATMDERRYTPRAHLD-----EGG 712
Db 691 -KMDVBLCFKVDSSAVFSEPLGRGMSLNFVTDVDTQKQRLQACADRSCCOSCLMKWSGG 749
QY 713 DRFTNRVALLSSGOBLERINFHVLDTADYVVKVPTFSVSESLRDPDH-----GPMLDGWM 767
Db 750 SSLCEHFLISTEBELCS-----DDCFSNITIKVSEYFQTSERRNHPNILDHYK 800
QY 768 PTLURSVPTFWNGCNEDEHCVPLDARSDDLPTAMEYCORVLKRPQDCSAYTLSDFTT 827
Db 801 EPSAIFQLPYEKDKCNKVFCAEI-----QLTTAIS-----QOD----- 834
QY 828 VFIIESTRORVAVEATLENGENAYSTVLTNSQSANLQFASLQKSDSDGSEICVNBERR 887
Db 835 --LVVGITKEVTMMNISLTNSGDSYNTMNALYPRNLQFKK-IQKPLSP-DIQCDPFPV 890
QY 888 LQKOV--CNVSYPPFRARAKVAFELDSFESKSTFLHHLLEIELAAGSDNERDSTKDNVA 945
Db 891 ASVLWAKCKIGHPIK-RSSVNVSVTWQLBESIPNRTADITVTISNKSRLARETH-- 947
QY 946 PLRFLHLYEA-----DVLFP--TRSSLSLHVEVLKNSSLERYDVGIGPPSPCFRIQNLGLF 998

```

|    |      |   |      |
|----|------|---|------|
| Db | 948  | SLOFRHAFI AVL SRPSVMYNTSQSSSDHKEFFNVHGENH-----FGAVFOLQ----- | 996  |
| Qy | 999  | PIHGMNKKIPIATRSNRLKLRDPLTDEANTSCNNGNSTEYRPTFVEEDLRAPOL      | 1058 |
| Db | 997  | -----ICVPITLRD-LQIRVKHLTKTQAHTEC-----TQSBPTCGSDPVQNVVEW     | 1042 |
| Qy | 1059 | NH-----SNSDVVSINCNIRLVPNQEINFHLLGNLWLRSLKALKY-----KSMKIMV   | 1105 |
| Db | 1043 | HSVICAITSNKENVTVAAEISMGHTKQL-----LRDISLQILGEISFNKSLYEG      | 1093 |
| Qy | 1106 | NAALQRFHSPFIIFREEDPSRQIEFEISKQEDWQVFIMIVGSTLGGLLALLLVIALRK  | 1165 |
| Db | 1094 | NAENHRTKITVIFLKEEKPH-----SLP--LIIGSSIGLLVLVVIILFK           | 1138 |
| Qy | 1166 | LGFERSARRRREPGLDPTPK  | 1185 |
| Db | 1139 | CGPPK--RKYQLNLESVRK   | 1156 |

Search completed: June 24, 2004, 17:55:41  
Job time : 64 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 17:53:06 ; Search time 64 Seconds  
(without alignments)  
5244.788 Million cell updates/sec

Title: US-09-980-403-2  
Perfect score: 6224  
Sequence: 1 MDLPRGLVVALSLWPGFT.....FRSARRRRPEGLDTPKYLE 1188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues  
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : A.Geneseq\_29Jan04:\*
- 1: Genesepi1980s:\*
  - 2: Genesepi1990s:\*
  - 3: Genesepi2000s:\*
  - 4: Genesepi2001s:\*
  - 5: Genesepi2002s:\*
  - 6: Genesepi2003as:\*
  - 7: Genesepi2003bs:\*
  - 8: Genesepi2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description        |
|------------|--------|-------------|--------|-------|--------------------|
| 1          | 6224   | 100.0       | 1188   | 4     | AAB30929 Amino aci |
| 2          | 6196   | 99.6        | 1188   | 4     | Aau14231 Human nov |
| 3          | 6196   | 99.6        | 1188   | 4     | Aab50085 Human A25 |
| 4          | 6196   | 99.6        | 1188   | 5     | Aau10551 Human A25 |
| 5          | 6192.5 | 99.5        | 1189   | 4     | Abg12949 Novel hum |
| 6          | 6192   | 99.5        | 1188   | 4     | Aau14467 Human nov |
| 7          | 6192   | 99.5        | 1188   | 7     | Ade09956 Novel pro |
| 8          | 6188.5 | 99.4        | 1189   | 3     | Abr25582 ITGA11 pr |
| 9          | 6188.5 | 99.4        | 1189   | 6     | Abr58364 Human NOV |
| 10         | 6188.5 | 99.4        | 1189   | 6     | Ada27054 Human nov |
| 11         | 6188.5 | 99.4        | 1189   | 7     | Ade63570 Human Pro |
| 12         | 6188.5 | 99.4        | 1189   | 8     | Ade86584 Novel hum |
| 13         | 5780   | 92.9        | 1120   | 6     | Abr58365 Human NOV |
| 14         | 5693   | 91.5        | 1188   | 4     | Aab50087 Murine A2 |
| 15         | 5693   | 91.5        | 1188   | 5     | Aau10552 Murine A2 |
| 16         | 5383   | 86.5        | 1034   | 3     | Aab25590 Protein e |
| 17         | 5383   | 86.5        | 1034   | 6     | Ada27062 Human nov |
| 18         | 5383   | 86.5        | 1034   | 8     | Ade86592 Novel hum |
| 19         | 3621   | 58.2        | 707    | 4     | Aau19663 Human nov |
| 20         | 3621   | 58.2        | 707    | 5     | Abp47883 Human pol |
| 21         | 3621   | 58.2        | 707    | 7     | Adc10845 Human ext |
| 22         | 3182   | 51.1        | 688    | 5     | Abb72300 Rat prote |
| 23         | 3113   | 50.0        | 696    | 5     | Abb72289 Rat prote |
| 24         | 2518   | 40.5        | 545    | 5     | Abb72288 Murine pr |
| 25         | 2403.5 | 38.6        | 1167   | 3     | Aay32242 Human int |

|    |        |      |      |   |          |                    |
|----|--------|------|------|---|----------|--------------------|
| 26 | 2401.5 | 38.5 | 1152 | 4 | AAB64657 | Aab64657 Human sec |
| 27 | 2375.5 | 38.2 | 1167 | 4 | AAB64584 | Aab64584 Human sec |
| 28 | 2375.5 | 38.2 | 1167 | 6 | ABP99490 | Abp99490 Human gen |
| 29 | 2375.5 | 38.2 | 1167 | 6 | ABR00964 | Abr00964 Human gen |
| 30 | 2375.5 | 38.2 | 1167 | 6 | ADA44026 | Ada44026 Human sec |
| 31 | 2373.5 | 38.1 | 1152 | 4 | AAB64658 | Aab64658 Human sec |
| 32 | 2319   | 37.3 | 1132 | 3 | AAY32243 | Aay32243 Human int |
| 33 | 2166.5 | 34.8 | 1179 | 5 | ABB90759 | Abb90759 Human Tum |
| 34 | 2166.5 | 34.8 | 1179 | 5 | ABP64915 | Abp64915 Human pro |
| 35 | 2166.5 | 34.8 | 1179 | 6 | ABUS4466 | Abu54466 Human int |
| 36 | 2159   | 34.7 | 1151 | 8 | ADE86652 | Ade86652 Human int |
| 37 | 2149.5 | 34.5 | 1180 | 5 | ABB90788 | Abb90788 Rat Tumou |
| 38 | 2149.5 | 34.5 | 1180 | 6 | ABUS4495 | Abu54495 Mouse tum |
| 39 | 2149.5 | 34.5 | 1180 | 7 | ADE63568 | Ade63568 Rat Prote |
| 40 | 2093   | 33.6 | 1183 | 2 | AAy07728 | Aay07728 Armenian  |
| 41 | 1863   | 29.9 | 1181 | 6 | ABU03548 | Abu03548 Angiogene |
| 42 | 1863   | 29.9 | 1181 | 6 | ABRS9703 | Abr59703 Human viA |
| 43 | 1863   | 29.9 | 1181 | 6 | AAG79775 | Aag79775 Alpha2 in |
| 44 | 1863   | 29.9 | 1181 | 6 | ABU03616 | Abu03616 Human exp |
| 45 | 1863   | 29.9 | 1181 | 6 | ABU03614 | Abu03614 Human exp |

ALIGNMENTS

RESULT 1  
AAB30929  
ID AAB30929 standard; protein; 1188 AA.

XX AAB30929;

AC AAB30929;

DT 02-APR-2001 (first entry)

XX Amino acid sequence of a human alpha11 integrin chain.

KW Human; integrin; alpha11 subunit; fibroblast; muscle cell; chondrocyte;

KW osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;

KW wound healing; trauma; rheumatoid arthritis; osteoarthritis;

KW osteoporosis; cartilage damage; bone damage; cartilage.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..22

FT Region /note= "signal peptide"

FT Domain /note= "leucine zipper"

FT Domain /note= "transmembrane domain"

XX WO2000075187-A1.

XX 14-DEC-2000.

XX 31-MAY-2000; 2000WO-SE001135.

XX 03-JUN-1999; 99SE-00002056.

XX (ACTI-) ACTIVE BIOTECH AB.

XX Gullberg D;

XX WPI; 2001-071061/08.

XX N-PSDB; AAC86871.

XX Integrin subunit alpha 11 or integrin heterodimer comprising subunit

PT alpha 11 in association with subunit beta, useful for treating muscle

PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.

XX Disclosure; Fig 2a-c; 79pp; English.

XX The present sequence represents a human integrin subunit, designated

CC alpha11. The alpha11 polynucleotide and polypeptide are useful as markers

CC of cell target molecules, such as fibroblasts, muscle cells,  
 CC chondrocytes, osteoblasts, mesenchymally derived cells or stem cells.  
 CC They are also used for determining the differential-stage of cells during  
 CC differentiation, development in pathological conditions, in tissue  
 CC regeneration, in transplantation or in therapeutic and physiological  
 CC repair of tissues. The pathological conditions involving subunit alpha1  
 CC are selected from damage of cells, muscle dystrophy, fibrosis, wound  
 CC healing, trauma, rheumatoid arthritis, osteoarthritis and osteoporosis,  
 CC damage of cartilage and bone, and cartilage and bone diseases. The  
 CC polypeptide is useful for detecting the formation of cartilage during  
 CC embryonic development, for detecting physiological therapeutic repair of  
 CC cartilage and muscle, for selection and analysis, or for sorting,  
 CC isolating or purification of chondrocytes and muscle cells, for detecting  
 CC regeneration of cartilage or chondrocytes during transplantation of  
 CC cartilage or chondrocytes during transplantation of cartilage or  
 CC chondrocytes, respectively, or of muscle or muscle cells during  
 CC transplantation of muscle or muscle cells, respectively, and for studies  
 CC of differentiation or chondrocytes or muscle cells  
 XX  
 SQ Sequence 1188 AA;  
 Query Match 100.0%; Score 6224; DB 4; Length 1188;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDLPRLGVWAWLSLWPGFTDFTNMDTRKRVTPGSRTPAFGTYVQCHDISGNKWLWVGA 60  
 DB 1 MDLPRLGVWAWLSLWPGFTDFTNMDTRKRVTPGSRTPAFGTYVQCHDISGNKWLWVGA 60  
 QY 61 PLEINGYQKTDGVYKCPVIGHNTKLNLRVTLVSNVSRKDNRLGLSLATNPKNSFLA 120  
 DB 61 PLEINGYQKTDGVYKCPVIGHNTKLNLRVTLVSNVSRKDNRLGLSLATNPKNSFLA 120  
 QY 121 CSPLWSHECGSSYTTGMCGRVNSNFRFSKTVAPALQRCQTYMDIIVLDGNSIYPWVE 180  
 DB 121 CSPLWSHECGSSYTTGMCGRVNSNFRFSKTVAPALQRCQTYMDIIVLDGNSIYPWVE 180  
 QY 181 VQHLINILKFFIIGPQIQGVVYQGEDVHVFHFLNDYRSVKDVEAASHIQRGTET 240  
 DB 181 VQHLINILKFFIIGPQIQGVVYQGEDVHVFHFLNDYRSVKDVEAASHIQRGTET 240  
 QY 241 RTAFGIETFAKSEAFQKGRGAKKMWIVITDGHSDSPDLKVIQOSRDNTVRYAVVL 300  
 DB 241 RTAFGIETFAKSEAFQKGRGAKKMWIVITDGHSDSPDLKVIQOSRDNTVRYAVVL 300  
 QY 301 GYNRRGINPETFLNEIKYASDDPKHFFNVDEAALKDIDALGRIFSLGNTKNET 360  
 DB 301 GYNRRGINPETFLNEIKYASDDPKHFFNVDEAALKDIDALGRIFSLGNTKNET 360  
 QY 361 SPGLEMSQTCFSSHEVEDGVLGAVGAYDNGAVLXKTSAGKVIPLRESYLKFPPELKN 420  
 DB 361 SPGLEMSQTCFSSHEVEDGVLGAVGAYDNGAVLXKTSAGKVIPLRESYLKFPPELKN 420  
 QY 421 HGAYLGVYTVTSVSSRGRVYVAGAPFNHTGKVLFTMHNESLATHQAMRGQQIGSYF 480  
 DB 421 HGAYLGVYTVTSVSSRGRVYVAGAPFNHTGKVLFTMHNESLATHQAMRGQQIGSYF 480  
 QY 481 GSEITTSVDIDGQVTDVLLVAGAPMYFNEGERGKVVYVELRQNFVYNGTLKDSHSYQNA 540  
 DB 481 GSEITTSVDIDGQVTDVLLVAGAPMYFNEGERGKVVYVELRQNFVYNGTLKDSHSYQNA 540  
 QY 541 RFGSSIASVRLDNQDSYNDVWVWVAGAPLDDNHAGAIYIPHGFRGSIILKTPQRIITASELATG 600  
 DB 541 RFGSSIASVRLDNQDSYNDVWVWVAGAPLDDNHAGAIYIPHGFRGSIILKTPQRIITASELATG 600  
 QY 601 LOYFGCSIHGQDLNEDGLDILAVGALGNVILWSRPVQINASLHPEPSKINIFHRDCK 660  
 DB 601 LOYFGCSIHGQDLNEDGLDILAVGALGNVILWSRPVQINASLHPEPSKINIFHRDCK 660  
 QY 661 RSGRDATCLAAFLCFTPIFLAPHQTTTGVIRINATMDERRYPTRAHLDSGGDRFTNRAV 720  
 DB 661 RSGRDATCLAAFLCFTPIFLAPHQTTTGVIRINATMDERRYPTRAHLDSGGDRFTNRAV 720

QY 721 LLSSGOELCERINFHVLDTADYVKPVTSPVSEYSLDDPDHGMPLDDGMPITLRSVPFFWNG 780  
 DB 721 LLSSGOELCERINFHVLDTADYVKPVTSPVSEYSLDDPDHGMPLDDGMPITLRSVPFFWNG 780  
 QY 781 CNEDEHCVPLDLVLDARSPLTAMBYCQVLRKPAQDCSAYTSLSPDTTFFIIESTRQYAV 840  
 DB 781 CNEDEHCVPLDLVLDARSPLTAMBYCQVLRKPAQDCSAYTSLSPDTTFFIIESTRQYAV 840  
 QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERLQKQVCNVSYPFF 900  
 DB 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERLQKQVCNVSYPFF 900  
 QY 901 RAKAKVAPRLDSBFSKSIPLHLEIELAAGSDSNERDSTKEDNVAPLAFHLKYEADVLT 960  
 DB 901 RAKAKVAPRLDSBFSKSIPLHLEIELAAGSDSNERDSTKEDNVAPLAFHLKYEADVLT 960  
 QY 961 RSSLSHVEVKLNSLERYDGIQPPSCIFRIQNLGLFPIHGMWMMKIPIPIATRSNRL 1020  
 DB 961 RSSLSHVEVKLNSLERYDGIQPPSCIFRIQNLGLFPIHGMWMMKIPIPIATRSNRL 1020  
 QY 1021 KLKDLTDEANTSCNINWNTSTYRTPVEEDLRAPQLNHSNSDVVSNINRLVPOEI 1080  
 DB 1021 KLKDLTDEANTSCNINWNTSTYRTPVEEDLRAPQLNHSNSDVVSNINRLVPOEI 1080  
 QY 1081 NFHLGNLWLSRLKALKYKSMKIMVNAALQRFHSPFIFREEDPSRQIEFEISKQEDQV 1140  
 DB 1081 NFHLGNLWLSRLKALKYKSMKIMVNAALQRFHSPFIFREEDPSRQIEFEISKQEDQV 1140  
 QY 1141 PIWIIIVGSTLGLLLELALLVLALRLKLGFRSARRRREGLDPTPKVLE 1188  
 DB 1141 PIWIIIVGSTLGLLLELALLVLALRLKLGFRSARRRREGLDPTPKVLE 1188  
 RESULT 2  
 AAU14231  
 ID AAU14231 standard; protein; 1188 AA.  
 XX  
 AC AAU14231;  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human novel protein #102.  
 XX  
 KW Human; novel protein; Antianemic; osteopathic; antiinflammatory;  
 KW immunomodulatory; cytostatic; neuroprotective; vulnary; nootropic;  
 KW antibacterial; antiallergic; dermatological; haemostatic; antisthmatic;  
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
 KW tissue regeneration; immune disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200155437-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 25-JAN-2001; 2001WO-US002623.  
 XX  
 PR 25-JAN-2000; 2000US-00491404.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 DR WPI; 2001-451939/48.  
 DR N-PSDB; AAS22536.  
 XX  
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,  
 PT nervous system disorders, and for regenerating bone and cartilage.  
 XX  
 PS Example 4; Page 578-581; 894pp; English.  
 XX

CC The invention relates to polynucleotides encoding novel human proteins or  
 CC their active domains. The polypeptides, polynucleotides and antibodies  
 CC raised against the polypeptides are used in a method of treatment of a  
 CC mammal and prevention of disorders caused by the aberrant protein  
 CC expression or activity. The polypeptides can be used as molecular weight  
 CC markers, food supplements, and in antibody production. The polypeptides  
 CC are used to identify compounds which bind to the polypeptides.  
 CC Polynucleotides of the invention are used as probes and primers, for  
 CC sequencing, for chromosome or gene mapping, in the production of  
 CC recombinant proteins, and in generating anti-sense DNA or RNA and in gene  
 CC therapy. Polypeptides of the invention can be used to target drugs to a  
 CC tumour, in assays to determine biological activity, to raise  
 CC antibodies/elicit an immune response, to determine quantitative protein  
 CC levels, as tissue markers, and to isolate receptors or ligands.  
 CC Polypeptides of the invention may also be useful in treating platelet  
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
 CC the proliferation, differentiation and survival of stem cells, as a  
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
 CC fungal infection or from autoimmunity, cancer, allergy, asthma, graft-  
 CC versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory  
 CC diseases, nervous system disorders, and infection. The present sequence  
 CC represents a protein of the invention  
 XX  
 SQ Sequence 1188 AA;

Query Match 99.6%; Score 61.96; DB 4; Length 1188;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 1184; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDLPRLGVAVWALSMPGDTDFNMTRKPRVPGSRVTAFFQVTVQOHDISGNKWLAVGA 60  
 DB 1 MDLPRLGVAVWALSMPGDTDFNMTRKPRVPGSRVTAFFQVTVQOHDISGNKWLAVGA 60

QY 61 PLETFYQKTYGVKCPVTHGCTKLNLRVTLNVSRKDNRLGLSLATNPKNDSPLA 120  
 DB 61 PLETFYQKTYGVKCPVTHGCTKLNLRVTLNVSRKDNRLGLSLATNPKNDSPLA 120

QY 121 CSPPLASHECGSSYYTTCMSRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSNYPWVE 180  
 DB 121 CSPPLASHECGSSYYTTCMSRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSNYPWVE 180

QY 181 VQHFILNLIKPKYIPGQIQVGVQGEDVWHEFHLNDYRSVKDWEAASHIEQGGTET 240  
 DB 181 VQHFILNLIKPKYIPGQIQVGVQGEDVWHEFHLNDYRSVKDWEAASHIEQGGTET 240

QY 241 RTAFGIEPARSAFQKGRKGAKKVMIVITDGESHSDPLEKVIQOQSERDNYRVAVVL 300  
 DB 241 RTAFGIEPARSAFQKGRKGAKKVMIVITDGESHSDPLEKVIQOQSERDNYRVAVVL 300

QY 301 GYNNRRGINPETFLEIKYIASDPDDKHFNFVTDAAALKDVIDALGDRIFSLGNTKNET 360  
 DB 301 GYNNRRGINPETFLEIKYIASDPDDKHFNFVTDAAALKDVIDALGDRIFSLGNTKNET 360

QY 361 SFGLMSQGRFSHVVDGVLGAGYDWNAGVLTSETSGAKVPIPLRESYLKEPPEELKN 420  
 DB 361 SFGLMSQGRFSHVVDGVLGAGYDWNAGVLTSETSGAKVPIPLRESYLKEPPEELKN 420

QY 421 HGAYLGTYVTSVVSSRQGRVYVAGAPRFNHTCKVILFTMHNRSITIHQAEQGOIGSYF 480  
 DB 421 HGAYLGTYVTSVVSSRQGRVYVAGAPRFNHTCKVILFTMHNRSITIHQAEQGOIGSYF 480

QY 481 GSEITSVDIDGQGVTDVLLVGAPMYPNEGRGKVIYVELRQNVFYNGTLKDSYSYQNA 540  
 DB 481 GSEITSVDIDGQGVTDVLLVGAPMYPNEGRGKVIYVELRQNVFYNGTLKDSYSYQNA 540

QY 541 RFGSSIASVRDLNQDSYNDVWVYCAPLEDNHAGAIYIFHGFRGSILKTPKORITASELATG 600  
 DB 541 RFGSSIASVRDLNQDSYNDVWVYCAPLEDNHAGAIYIFHGFRGSILKTPKORITASELATG 600

QY 601 LQYFGCSITHGQDLNEDGLIDLAVALGNVILWSRPVVQINASLHFEPSSKINIFHRDCK 660

DB 602 LQYFGCSITHGQDLNEDGLIDLAVALGNVILWSRPVVQINASLHFEPSSKINIFHRDCK 660  
 QY 661 RSGRDATCLAAFLCFTTIFLAPHFQTTTIGIRYNATMDERYTTPRAHLDEGDRFTNRAV 720  
 DB 661 RSGRDATCLAAFLCFTTIFLAPHFQTTTIGIRYNATMDERYTTPRAHLDEGDRFTNRAV 720  
 QY 721 LLSGGQELCRINEHVLDTADYKVPVTFSEVSEYLEDPDHGPMLDGDPMTTLRVSPVWNG 780  
 DB 721 LLSGGQELCRINEHVLDTADYKVPVTFSEVSEYLEDPDHGPMLDGDPMTTLRVSPVWNG 780  
 QY 781 CNEDEHCVPLVDLARSGLPTAMEYQORVLKPAQDCSAVTLSEDTTVPFIIESTRQRVAV 840  
 DB 781 CNEDEHCVPLVDLARSGLPTAMEYQORVLKPAQDCSAVTLSEDTTVPFIIESTRQRVAV 840  
 QY 841 EATLENGENAYSTVNLISQSANIQPASLQKEDSDGSEIECVNEERLQKQVCNVSPFF 900  
 DB 841 EATLENGENAYSTVNLISQSANIQPASLQKEDSDGSEIECVNEERLQKQVCNVSPFF 900  
 QY 901 RAKAKVAFRLDSEFSKSIPLHLHLELAAGSDSNRSDTKEDNVAPLRFHLKYEADVLT 960  
 DB 901 RAKAKVAFRLDSEFSKSIPLHLHLELAAGSDSNRSDTKEDNVAPLRFHLKYEADVLT 960  
 QY 961 RSSLSHYVVKNSLSERYDGIQPPFSCIFRIQNLGLPFIHGMMKTIPIATRSNRL 1020  
 DB 961 RSSLSHYVVKNSLSERYDGIQPPFSCIFRIQNLGLPFIHGMMKTIPIATRSNRL 1020  
 QY 1021 KLRFDLDEANTSCNIWGNSTERYPTVEEDLRAPQLNHSNSDVWSINCIRLVPNOEI 1080  
 DB 1021 KLRFDLDEANTSCNIWGNSTERYPTVEEDLRAPQLNHSNSDVWSINCIRLVPNOEI 1080  
 QY 1081 NFHLGLMLWLSLKALKYKSMKIMVNAALQRFHSPIFREDPSRQIEFISQEDNQV 1140  
 DB 1081 NFHLGLMLWLSLKALKYKSMKIMVNAALQRFHSPIFREDPSRQIEFISQEDNQV 1140  
 QY 1141 PIIIVGSTIGLLALLLALLLALRKLGPFERSARRRPPGLDTPKVL 1188  
 DB 1141 PIIIVGSTIGLLALLLALLLALWKLGPFERSARRRPPGLDTPKVL 1188

RESULT 3  
 AAB50085  
 ID AAB50085 standard; protein; 1188 AA.  
 XX AAB50085;  
 AC AAB50085;  
 DT 19-MAR-2001 (first entry)  
 XX Human A259.  
 DE Human; A259; integrin alpha subunit; fibrosis; liver disease; lung;  
 XX kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;  
 XX rheumatoid arthritis.  
 XX Homo sapiens.  
 OS  
 PH Key  
 FT Domain  
 FT Peptide  
 FT Protein  
 FT Protein  
 FT Domain  
 FT Domain  
 FT Domain  
 FT Domain  
 FT Domain  
 FT Domain  
 FT Domain

Location/Qualifiers  
 1. .1141  
 /label= Extracellular\_domain  
 1. .22  
 /label= Signal\_peptide  
 23. .1188  
 /label= Mature\_protein  
 39. .74  
 /label= Integrin\_alphasubunit\_repeat\_domain\_#1  
 115. .157  
 /label= Integrin\_alphasubunit\_repeat\_domain\_#2  
 164. .345  
 /label= I\_domain  
 367. .392  
 /label= Integrin\_alphasubunit\_repeat\_domain\_#3  
 421. .455  
 /label= Integrin\_alphasubunit\_repeat\_domain\_#4

FT Domain 478..516  
 FT /label= Integrin\_alpha subunit\_repeat\_domain\_#5  
 FT Domain 540..575  
 FT /label= Integrin\_alpha subunit\_repeat\_domain\_#6  
 FT Domain 602..640  
 FT /label= Integrin\_alpha subunit\_repeat\_domain\_#7  
 FT Domain 1142..1164  
 FT /label= Transmembrane\_domain  
 FT Domain 1165..1188  
 FT /label= Cytoplasmic\_domain  
 XX W0200073339-AL.  
 XX 07-DEC-2000.  
 XX 15-MAY-2000; 2000MO-US013262.  
 XX 28-MAY-1999; 99US-00322790.  
 PR 27-APR-2000; 2000US-00561263.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA Pan Y, Lora JM;  
 PI WPI: 2001-041142/05.  
 DR N-PSDB; AAC91901; AAC91902.  
 XX Nucleic acid encoding alpha-integrin subunits, useful for treatment and  
 PT diagnosis of fibrosis, e.g. of the liver.  
 XX Claim 8; Fig 1; 164pp; English.  
 CC The present sequence is human integrin alpha subunit, A259. A259 is  
 CC homologous with the alpha1 and alpha10 integrin subunits and is  
 CC overexpressed in fibrosis. A259 is implicated in regulation of  
 CC proliferation, differentiation and/or function of many different cell  
 CC types. Inhibitors of A259 activity are useful for the treatment of liver  
 CC disease, particularly fibrosis, and also fibrosis in other organs  
 CC (specifically lung and kidney). In addition, A259 can be used for  
 CC treatment and prevention of cancer, osteoporosis, acute myeloid  
 CC leukaemia, HIV infection, and rheumatoid arthritis  
 XX Sequence 1188 AA;  
 SQ  
 Query Match 99.6%; Score 6196; DB 4; Length 1188;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 1184; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MDLPRLGVAVALS LWPFGFTDTFNM DTRKPRVPGSRTAPFGYTVQOHDISGNKMLVWGA 60  
 DB 1 MDLPRLGVAVALS LWPFGFTDTFNM DTRKPRVPGSRTAPFGYTVQOHDISGNKMLVWGA 60  
 QY 61 PLETNGYQKTDYVKCPVHNGCTKLN LGRTLSNVSRKDNKRLGLSLATNPXNSFLA 120  
 DB 61 PLETNGYQKTDYVKCPVHNGCTKLN LGRTLSNVSRKDNKRLGLSLATNPXNSFLA 120  
 QY 121 CSPLWHECGSSYTTGMC SRVNSNFRFSKTVAPALQRCQTYMDIIVLDGNSNIYPWVE 180  
 DB 121 CSPLWHECGSSYTTGMC SRVNSNFRFSKTVAPALQRCQTYMDIIVLDGNSNIYPWVE 180  
 QY 181 VQHFLNLLKFFIGQIQGVGVQGEDVHVFHNDYRSVNDVVEAASHIQRGTET 240  
 DB 181 VQHFLNLLKFFIGQIQGVGVQGEDVHVFHNDYRSVNDVVEAASHIQRGTET 240  
 QY 241 RTAFGIEFARSEAFQGRGAKKVMIVITDGSHPDPLEKVIQOSERDNVTRYAVVL 300  
 DB 241 RTAFGIEFARSEAFQGRGAKKVMIVITDGSHPDPLEKVIQOSERDNVTRYAVVL 300  
 QY 301 GYNRGINPETHLEIKYIASPDQKPNVNDVDEALKDVIDALGDRIFESLEGTNKNET 360  
 DB 301 GYNRGINPETHLEIKYIASPDQKPNVNDVDEALKDVIDALGDRIFESLEGTNKNET 360  
 QY 361 SFGLEMSQTGFSSHVVEDGVLLGAVGYDNGAVLKETSAGKVIPLRESYLKEFPEELKN 420

DB 361 SFGLEMSQTGFSSHVVEDGVLLGAVGYDNGAVLKETSAGKVIPLRESYLKEFPEELKN 420  
 QY 421 HGAYLGYTVTSVVSRRQGRVYVAGAPRNFNTGKVLFTWNNRSLTIHQAGGQQIGSYF 480  
 DB 421 HGAYLGYTVTSVVSRRQGRVYVAGAPRNFNTGKVLFTWNNRSLTIHQAGGQQIGSYF 480  
 QY 481 GSEITSDVDGSDVTVLLVAGAPMYFNEGREGKGVYVYELQRNRRFYANGTLKDSHSYQNA 540  
 DB 481 GSEITSDVDGSDVTVLLVAGAPMYFNEGREGKGVYVYELQRNRRFYANGTLKDSHSYQNA 540  
 QY 541 RGSSTIASVRDLNQDSYNDVVGAPLEONHAGAIYIFHGFRGSIILKTPKORITASELATG 600  
 DB 541 RGSSTIASVRDLNQDSYNDVVGAPLEONHAGAIYIFHGFRGSIILKTPKORITASELATG 600  
 QY 601 LQYFGCSIHQDLNEDGLIDLAVGALGNVILNSRPVQINASLHFPSPKINI FHRDCK 660  
 DB 601 LQYFGCSIHQDLNEDGLIDLAVGALGNVILNSRPVQINASLHFPSPKINI FHRDCK 660  
 QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTTPRAHLDGGRFTNRAV 720  
 DB 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTTPRAHLDGGRFTNRAV 720  
 QY 721 LLSQQLCERINFHVLDTADYVKPVTSEVYSLEDDPDHGMKLDGMPPTLRSVFPFNG 780  
 DB 721 LLSQQLCERINFHVLDTADYVKPVTSEVYSLEDDPDHGMKLDGMPPTLRSVFPFNG 780  
 QY 781 CNEDEHCVPDLVDARSDLPTAMEYCORVLRKPAQDCSAYTLSPDTTFIESTQRVAV 840  
 DB 781 CNEDEHCVPDLVDARSDLPTAMEYCORVLRKPAQDCSAYTLSPDTTFIESTQRVAV 840  
 QY 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGSI ECNWEERRKQKQCNVSYPPF 900  
 DB 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGSI ECNWEERRKQKQCNVSYPPF 900  
 QY 901 RAKAVAPRLDSEFSKSIPLHLHLEISLAAGSDNSNERDSTKEDNVAPLPHLKYEADVLPT 960  
 DB 901 RAKAVAPRLDSEFSKSIPLHLHLEISLAAGSDNSNERDSTKEDNVAPLPHLKYEADVLPT 960  
 QY 961 RSSLSHVEVKLNSLERYDGI GPPFCIFRIONLGLPIHGMKMIITPIATRSNRL 1020  
 DB 961 RSSLSHVEVKLNSLERYDGI GPPFCIFRIONLGLPIHGMKMIITPIATRSNRL 1020  
 QY 1021 KLRFPLTDEANTSCNMGNSNTEYRPTVEEDLRAPQLNHSNSDVVSINCMLRPNQEI 1080  
 DB 1021 KLRFPLTDEANTSCNMGNSNTEYRPTVEEDLRAPQLNHSNSDVVSINCMLRPNQEI 1080  
 QY 1081 NFHLGNLWLSLKLKLYSKMKIMVNAALQRFSPFI FREDPSRQIEFETSKQEDWQV 1140  
 DB 1081 NFHLGNLWLSLKLKLYSKMKIMVNAALQRFSPFI FREDPSRQIEFETSKQEDWQV 1140  
 QY 1141 PIWIIIVGSTLGGELLALLLVALLALRKLGFPSARRRREFGLDPTPKVLE 1188  
 DB 1141 PIWIIIVGSTLGGELLALLLVALLALRKLGFPSARRRREFGLDPTPKVLE 1188  
 RESULT 4  
 AAU10551  
 ID AAU10551 standard; protein; 1188 AA.  
 XX  
 AC AAU10551;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Human A259 polypeptide.  
 XX  
 DE Human; A259; integrin alpha subunit; integrin alpha 10; secreted protein;  
 KW liver disease; fibrosis; lung; kidney; bone associated disorder; blood;  
 KW cartilage associated disorder; haematopoietic disorder; bone marrow;  
 KW immune related disease; apoptotic disorder; neuronal tissue disease;  
 KW neurodegenerative disease; gene therapy; cancer; cytotoxic; osteopathic;  
 KW nephrotropic; immunomodulator; anti-inflammatory; neuroprotective;  
 KW antiarthritic; antianaemic; antiallergic; antiasthmatic; dermatological;



antidiabetic; anticonvulsant; antiparkinsonian.

Homo sapiens.

Key Location/Qualifiers

1..1141 /note= "Extracellular domain"

1..22 /note= "Signal peptide"

23..1188 /note= "Mature human A259"

37..90 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"

115..157 /note= "Integrin alpha repeat domain"

164..345 /note= "I domain or Von Willebrand Factor type A domain"

367..392 /note= "Integrin alpha repeat domain"

421..472 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"

476..532 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"

538..593 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"

600..654 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"

1142..1164 /note= "Transmembrane domain"

1165..1188 /note= "Cytoplasmic domain"

WC200181414-A2.

01-NOV-2001.

27-APR-2001; 2001WO-05013516.

27-APR-2000; 2000US-00561263.

(MILL-) MILLENNIUM PHARM INC.

Pan Y, Lora J;

WPI; 2002-041397/05.

N-PSDB; AAS16873.

New A259 nucleic acids and polypeptides, which comprise integrin alpha subunit, useful for diagnosing, preventing or treating e.g. liver disease, kidney or lung fibrosis, cancers, blood disorders or immune related diseases.

Claim 9; Fig 1; 168pp; English.

The invention relates to human and murine A259 nucleic acid molecules which encode secreted proteins with homology to integrin alpha subunits. Specifically to integrin alpha 10. The A259 polypeptide and nucleic acid are useful for treating liver disease or fibrosis, particularly kidney fibrosis or lung fibrosis. The A259 polypeptide and nucleic acid are also useful for diagnosing, preventing or treating cartilage and bone associated disorders (such as bone cancer, achondroplasia, myeloma, fibrous dysplasia, scoliosis, osteoarthritis, osteosarcoma and osteoporosis), bone marrow, blood and haematopoietic disorders (such as acute myeloid leukaemia, haemophilia, anaemia and thalassemia), immune related diseases (such as HIV, viral infections, cancers, T cell autoimmune disorders e.g. AIDS, and allergic inflammatory disorders e.g. asthma and psoriasis), apoptotic disorders (such as systemic lupus erythematosus and insulin-dependent diabetes mellitus), diseases of the

CC neuronal tissues (such as epilepsy and muscular dystrophy) and  
CC neurodegenerative diseases (such as Parkinson's disease and Huntington's  
CC disease). This sequence represents the human A259 polypeptide  
XX  
SQ Sequence 1188 AA;

Query Match 99.6%; Score 6196; DB 5; Length 1188;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1184; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDLPRGLVAVALSILMPGFTDTFNMDTRKPRVIGSRRTAFPGYTVQOHDISGNKWLTVGA 60  
DB 1 MDLPRGLVAVALSILMPGFTDTFNMDTRKPRVIGSRRTAFPGYTVQOHDISGNKWLTVGA 60  
QY 61 PLENGVQKTGDVYKCPVIHGNCTKLANGRVTLNVSERKDNMRIGLSLATNPKDNSPLA 120  
DB 61 PLENGVQKTGDVYKCPVIHGNCTKLANGRVTLNVSERKDNMRIGLSLATNPKDNSPLA 120  
QY 121 CSPLWSHECGSSYITGMCNRSNFRSKTAPALQCTYMDIVIVLDSNSIYPWVE 180  
DB 121 CSPLWSHECGSSYITGMCNRSNFRSKTAPALQCTYMDIVIVLDSNSIYPWVE 180  
QY 181 VQFLINILKXFFYIGPGQIQGVVQYGEDVVHFEHLNDYRSKDVVEAASHIEQGGTET 240  
DB 181 VQFLINILKXFFYIGPGQIQGVVQYGEDVVHFEHLNDYRSKDVVEAASHIEQGGTET 240  
QY 241 RTAGIEPARSEATQKGRKCAKVMIVITDGRSHSDPLEKVIQOOSERDNRVAVAVL 300  
DB 241 RTAGIEPARSEATQKGRKCAKVMIVITDGRSHSDPLEKVIQOOSERDNRVAVAVL 300  
QY 301 GYNNRRGINPETFLNEIKYIASDDPDKHFFNVTDAAALKDIDVALGDRIFSLGNTNKNET 360  
DB 301 GYNNRRGINPETFLNEIKYIASDDPDKHFFNVTDAAALKDIDVALGDRIFSLGNTNKNET 360  
QY 361 SFGLEMSQTFSSHVVDGVLIGAVYDNGAVLKETSAGKVIPLRESYLKEPFEELKN 420  
DB 361 SFGLEMSQTFSSHVVDGVLIGAVYDNGAVLKETSAGKVIPLRESYLKEPFEELKN 420  
QY 421 HGAVLGTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNNRSLTIHQMRGQOIGSYF 480  
DB 421 HGAVLGTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNNRSLTIHQMRGQOIGSYF 480  
QY 481 GSEITSDIDGVDVLLVAGPMYFNEGRGRGVYVYELQNRFPVYNGTKDKSHSYQNA 540  
DB 481 GSEITSDIDGVDVLLVAGPMYFNEGRGRGVYVYELQNRFPVYNGTKDKSHSYQNA 540  
QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGALVIPHGRGSLKTKPKORITASELATG 600  
DB 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGALVIPHGRGSLKTKPKORITASELATG 600  
QY 601 LQYFCCSIHQQLDNEGLDILAVGALGNVILMSRPVQINASLHPSPSKINIFHRDCK 660  
DB 601 LQYFCCSIHQQLDNEGLDILAVGALGNVILMSRPVQINASLHPSPSKINIFHRDCK 660  
QY 661 RSGRDATCLAAFLCFTRIFLAPHQTTTGTIRYNATMDERYTPRAHLDEGGDRTNRAV 720  
DB 661 RSGRDATCLAAFLCFTRIFLAPHQTTTGTIRYNATMDERYTPRAHLDEGGDRTNRAV 720  
QY 721 LLSGQQLCERINFHVLDTDADYVKPVPFSVEYSLEDDPHGPMDDGWPTTLRSVPFWNG 780  
DB 721 LLSGQQLCERINFHVLDTDADYVKPVPFSVEYSLEDDPHGPMDDGWPTTLRSVPFWNG 780  
QY 781 CNEDEHCVDPDILVLDARSPLTAMEYQORVLRKPAQDCSAYTILSDTTVTTFIESTQRVAV 840  
DB 781 CNEDEHCVDPDILVLDARSPLTAMEYQORVLRKPAQDCSAYTILSDTTVTTFIESTQRVAV 840  
QY 841 EATLENRGENTVTLNISQSANLQFASLIQKESDGSIECVNERERKQVCNVSYPFF 900  
DB 841 EATLENRGENTVTLNISQSANLQFASLIQKESDGSIECVNERERKQVCNVSYPFF 900  
QY 901 RAKAKVAPRLDSEFSKSIPLHLEIELAAGSDNSNERDSTKEDNVAPLPHIKYEADVLT 960  
DB 901 RAKAKVAPRLDSEFSKSIPLHLEIELAAGSDNSNERDSTKEDNVAPLPHIKYEADVLT 960



QY 961 RSSLSHYEVLNSSLERYDGIQPPSCIPRIOMGLFPIHGMWKITTIPIATRSNRL 1020  
Db 961 RSSLSHYEVLNSSLERYDGIQPPSCIPRIOMGLFPIHGMWKITTIPIATRSNRL 1020  
QY 1021 KLKDFLTDEANTSCNIGWSTVRPTPVEEDLRAPQLNHSNDDVVSINCNRLVNPQSI 1080  
Db 1021 KLKDFLTDEANTSCNIGWSTVRPTPVEEDLRAPQLNHSNDDVVSINCNRLVNPQSI 1080  
QY 1081 NFHLGNLWLSRLKALKYKSMKIMVNAALQORQFHSPPIFREEDPSQIFSEISKQEDWQV 1140  
Db 1081 NFHLGNLWLSRLKALKYKSMKIMVNAALQORQFHSPPIFREEDPSQIFSEISKQEDWQV 1140  
QY 1141 PIWIIWGSTLGLLALLVIALKLGFFPSARRRRPPGLDPTPKVLE 1188  
Db 1141 PIWIIWGSTLGLLALLVIALKLGFFPSARRRRPPGLDPTPKVLE 1188

RESULT 5  
ABG12949  
ID ABG12949 standard; protein; 1189 AA.  
XX AC ABG12949;  
XX DT 13-FEB-2002 (first entry)  
XX DE Novel human diagnostic protein #12940.  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX OS Homo sapiens.  
XX PN WO200175067-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US008631.  
XX PR 31-MAR-2000; 2000US-00540217.  
XX PR 23-AUG-2000; 2000US-00649167.  
XX PA (HVS-) HVSQ INC.  
XX PI Dmanac RT, Liu C, Tang YT;  
XX WPI; 2001-639362/73.  
XX DR N-PSDB; AAS77136.  
XX PT New isolated polynucleotide and encoded polypeptides, useful in  
XX PT diagnostics, forensics, gene mapping, identification of mutations  
XX PT responsible for genetic disorders or other traits and to assess  
XX PT biodiversity.  
XX PS Claim 20; SEQ ID NO 43308; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
XX and in recombinant production of (II). The polynucleotides are also used  
XX in diagnostics as expressed sequence tags for identifying expressed  
XX genes. (I) is useful in gene therapy techniques to restore normal  
XX activity of (II) or to treat disease states involving (II). (II) is  
XX useful for generating antibodies against it, detecting or quantitating a  
XX polypeptide in tissue, as molecular weight markers and as a food  
XX supplement. (II) and its binding partners are useful in medical imaging  
XX of sites expressing (II). (I) and (II) are useful for treating disorders  
XX involving aberrant protein expression or biological activity. The  
XX polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 1189 AA;

Query Match 99.5%; Score 6192.5; DB 4; Length 1189;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1185; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
QY 1 MDLPRGLVAVALSLLWPGFTDTFNNMTKPRVIFGSRTPAFGYTVQOHDISGNKWLTVGA 60  
Db 1 MDLPRGLVAVALSLLWPGFTDTFNNMTKPRVIFGSRTPAFGYTVQOHDISGNKWLTVGA 60  
QY 61 PLEWNGYOKTGDVYKCPVINGNCTKLNLRVTLNVSEKDNMRGLSLATNPKDNFLA 120  
Db 61 PLEWNGYOKTGDVYKCPVINGNCTKLNLRVTLNVSEKDNMRGLSLATNPKDNFLA 120  
QY 121 CSPLWSEHCSSYYTTGMCVRVNSNFRFSTKVPALQRCQTYMDIVIVLDGSNSIYPWVE 180  
Db 121 CSPLWSEHCSSYYTTGMCVRVNSNFRFSTKVPALQRCQTYMDIVIVLDGSNSIYPWVE 180  
QY 181 VQHFLINILKKFYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDVVEASHIEQGGTET 240  
Db 181 VQHFLINILKKFYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDVVEASHIEQGGTET 240  
QY 241 RTAFGIEFARSEAFQKGGKGAKKVMIVITDSEHSDSPLEKVIQSSERDNTVRYAVAVL 300  
Db 241 RTAFGIEFARSEAFQKGGKGAKKVMIVITDSEHSDSPLEKVIQSSERDNTVRYAVAVL 300  
QY 301 GYNNRGINPETFLNEIKYIASDDDDGHPFNVTDAALKDIDALGDRIFSLGNTKNMET 360  
Db 301 GYNNRGINPETFLNEIKYIASDDDDGHPFNVTDAALKDIDALGDRIFSLGNTKNMET 360  
QY 361 SFGLEMSQTFSSHVVVEDGVLGAVGYDWMGAVLKETSAGKVIPLRSYSLKEPPEELKN 420  
Db 361 SFGLEMSQTFSSHVVVEDGVLGAVGYDWMGAVLKETSAGKVIPLRSYSLKEPPEELKN 420  
QY 421 HGAYLGYTVTSVSSRQGRVTVAGAPRNFHTKVLFTMHNNRSLTIHQMRGQOIGSYF 480  
Db 421 HGAYLGYTVTSVSSRQGRVTVAGAPRNFHTKVLFTMHNNRSLTIHQMRGQOIGSYF 480  
QY 481 GSEITSVDIDGVDGTVLLVAGPMYFNEGRGKVVYVELRQNRVYNGTLKDSYQNA 540  
Db 481 GSEITSVDIDGVDGTVLLVAGPMYFNEGRGKVVYVELRQNRVYNGTLKDSYQNA 540  
QY 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGFRGSILKTPKQITASELATG 600  
Db 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGFRGSILKTPKQITASELATG 600  
QY 601 LQYFGCSITHGOLDNEDGLIDLAVGALGNVILWSRPVQINASHLFFPSKINIPIHRDCK 660  
Db 601 LQYFGCSITHGOLDNEDGLIDLAVGALGNVILWSRPVQINASHLFFPSKINIPIHRDCK 660  
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720  
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720  
QY 721 LLSQQRCERINFLVLOTADYKVTSEVSEYLEDPDHGMKLDGMPPTLRSVYPPFNG 780  
Db 721 LLSQQRCERINFLVLOTADYKVTSEVSEYLEDPDHGMKLDGMPPTLRSVYPPFNG 780  
QY 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTILSFDITVFIIESTQRVAV 840  
Db 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTILSFDITVFIIESTQRVAV 840  
QY 841 EATLENRENAYSTVLNISQSANLOPASLIQKESDGSIECNEERRIKQVCNVSYPFF 900  
Db 841 EATLENRENAYSTVLNISQSANLOPASLIQKESDGSIECNEERRIKQVCNVSYPFF 900  
QY 901 YAKAKVAFRLDSEPSKSIPLHLEITELAAAGSDNSNERDSTKEDNVAPLPHKYEADVLFT 960  
Db 901 YAKAKVAFRLDSEPSKSIPLHLEITELAAAGSDNSNERDSTKEDNVAPLPHKYEADVLFT 960

Db 901 RAKAVARLDFEFSKIFLHLLEIELAGSDNSNEDSTKENDVAPLRLHYEADVLT 960  
 Qy 961 RSSLSHYEVKLNSSLYRVDGIPPPSCIFRIQNLGLFPIHGMXMKITPIATRSNELL 1020  
 Db 961 RSSLSHYEVKLNSSLYRVDGIPPPSCIFRIQNLGLFPIHGMXMKITPIATRSNELL 1020  
 Qy 1021 KLRDLFTDB-ANTSCNINCNSTYRPTVREDLRAPOLNHSNDVVSINCNIRLVPKOE 1079  
 Db 1021 KLRDLFTDB-ANTSCNINCNSTYRPTVREDLRAPOLNHSNDVVSINCNIRLVPKOE 1080  
 Qy 1080 INFHLGNLWLSLXALKYKSMKIMVNAALQRFHSPFIIFREEDPSROIETFSIKQEDWQ 1139  
 Db 1081 INFHLGNLWLSLXALKYKSMKIMVNAALQRFHSPFIIFREEDPSROIETFSIKQEDWQ 1140  
 Qy 1140 VPIWIIIVGSTGLLHLLLVLLALRLKGLFFRSARREPERGLDPTPKVLE 1188  
 Db 1141 VPIWIIIVSTGLGLLLLVLLALWKLGLFFRSARREPERGLDPTPKVLE 1189

RESULT 6  
 AAU14467  
 ID AAU14467 standard; protein; 1188 AA.  
 XX AAU14467;  
 AC AAU14467;  
 DT 24-OCT-2001 (first entry)  
 XX Human novel protein #338.  
 DE Human; novel protein; Antianaemic; osteopathic; antiinflammatory;  
 KW immunomodulatory; cycostatic; neuroprotective; vulnerary; nootropic;  
 KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;  
 KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;  
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
 KW tissue regeneration; immune disorder.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 XX WO200155437-A2.  
 PN 25-JAN-2001; 2001WO-US002623.  
 XX 02-AUG-2001.  
 PD 25-JAN-2001; 2000US-00491404.  
 PR (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Drmanac RT;  
 PI WPI; 2001-451939/48.  
 XX N-PSDB; AAS22772.  
 DR Isolated polypeptides useful for treating anti-inflammatory diseases,  
 XX nervous system disorders, and for regenerating bone and cartilage.  
 PS Example 4; Page 828-831; 894pp; English.  
 XX The invention relates to polynucleotides encoding novel human proteins or  
 CC their active domains. The polypeptides, polynucleotides and antibodies  
 CC raised against the polypeptides are used in a method of treatment of a  
 CC mammal and prevention of disorders caused by the aberrant protein  
 CC expression or activity. The polypeptides can be used as molecular weight  
 CC markers, food supplements, and in antibody production. The polypeptides  
 CC are used to identify compounds which bind to the polypeptides.  
 CC Polynucleotides of the invention are used as probes and primers, for  
 CC sequencing, for chromosome or gene mapping, in the production of  
 CC recombinant proteins, and in generating anti-sense DNA or RNA and in gene  
 CC therapy. Polypeptides of the invention can be used to target drugs to a  
 CC tumour, in assays to determine biological activity, to raise  
 CC antibodies/elicit an immune response, to determine quantitative protein  
 CC levels, as tissue markers, and to isolate receptors or ligands.

CC Polypeptides of the invention may also be useful in treating platelet  
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
 CC the proliferation, differentiation and survival of stem cells, as a  
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
 CC fungal infection or from autoimmunity, cancer, allergy, asthma, graft-  
 CC versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory  
 CC diseases, nervous system disorders, and infection. The present sequence  
 CC represents a protein of the invention  
 XX  
 SQ Sequence 1188 AA;  
 Query Match 99.5%; Score 6192; DS 4; Length 1188;  
 Best Local Similarity 99.6%; Pred. NO. 0;  
 Matches 1183; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 MDLPRGLVAVALSILWPGFTDTFNMDTRKPRVIPSRTAPFGYTVQQHDSGNKWLVVGA 60  
 Db 1 MDLPRGLVAVALSILWPGFTDTFNMDTRKPRVIPSRTAPFGYTVQQHDSGNKWLVVGA 60  
 Qy 61 PLETNGYQKTGDVYKCPVIHGNCTKMLNLRVTLSNVSEKDNMRLGLSLATNPXNSFLA 120  
 Db 61 PLETNGYQKTGDVYKCPVIHGNCTKMLNLRVTLSNVSEKDNMRLGLSLATNPXNSFLA 120  
 Qy 121 CSPLSHECGSSVYTTGCMKSRVNSNFRFSKTVAPALQRCOTYMDIVIVLDSGNSIYPWVE 180  
 Db 121 CSPLSHECGSSVYTTGCMKSRVNSNFRFSKTVAPALQRCOTYMDIVIVLDSGNSIYPWVE 180  
 Qy 181 VQFLINILKFKYTGPOQIQGVVQYGEDVVBEPHLDYRSVKDVVEAASHIEQGGTET 240  
 Db 181 VQFLINILKFKYTGPOQIQGVVQYGEDVVBEPHLDYRSVKDVVEAASHIEQGGTET 240  
 Qy 241 RTAFGIHFASBEAFQKGRGAKKMWIIVTDGSHSDPDLKVKIQQSERDNVTYAVAVL 300  
 Db 241 RTAFGIHFASBEAFQKGRGAKKMWIIVTDGSHSDPDLKVKIQQSERDNVTYAVAVL 300  
 Qy 301 GYNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSECTNKNET 360  
 Db 301 GYNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSECTNKNET 360  
 Qy 361 SFGLEMSQTGFSHVVEDGVLLGNVGYDNGAVLAKETSAGKVIPLRESYLKEPPELKN 420  
 Db 361 SFGLEMSQTGFSHVVEDGVLLGNVGYDNGAVLAKETSAGKVIPLRESYLKEPPELKN 420  
 Qy 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFETGKVIPLFMENRSLTHQAMRGQIGSYF 480  
 Db 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFETGKVIPLFMENRSLTHQAMRGQIGSYF 480  
 Qy 481 GSEITSVDIDGDGVTDLVLLVAGAPMYFNEGRERKQVYVELRQNFVYNGTLKDSHSYQNA 540  
 Db 481 GSEITSVDIDGDGVTDLVLLVAGAPMYFNEGRERKQVYVELRQNFVYNGTLKDSHSYQNA 540  
 Qy 541 RFGSSIASVRDLNODSVNDVVGAPLEDNHAGAIYIPHGFRGSLTKPKQITASELATG 600  
 Db 541 RFGSSIASVRDLNODSVNDVVGAPLEDNHAGAIYIPHGFRGSLTKPKQITASELATG 600  
 Qy 601 LQYFGCSIHQOLDNLBDGLIDLAVGALGNVILWSPVVOINASLHFPSPKINIIFHDCK 660  
 Db 601 LQYFGCSIHQOLDNLBDGLIDLAVGALGNVILWSPVVOINASLHFPSPKINIIFHDCK 660  
 Qy 661 RSRGDATCLAAFLCFTPIPLAHPFQTTVGIRYNATMDERRYTPRAHLDGGDRFTNRVAV 720  
 Db 661 RSRGDATCLAAFLCFTPIPLAHPFQTTVGIRYNATMDERRYTPRAHLDGGDRFTNRVAV 720  
 Qy 721 LLSGGQELCERINPHVLDATYVYKPTVFSVEYSLEDDPDHGMPLDDGPTTLRVSPFWNG 780  
 Db 721 LLSGGQELCERINPHVLDATYVYKPTVFSVEYSLEDDPDHGMPLDDGPTTLRVSPFWNG 780  
 Qy 781 CNEDEHCVPDLVLDARSDDLPTAMEYCORVLKPAQDCSAVTLSDFTVFTIESTRORVAV 840  
 Db 781 CNEDEHCVPDLVLDARSDDLPTAMEYCORVLKPAQDCSAVTLSDFTVFTIESTRORVAV 840

QY 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNERRLQKQVCNVSYPFF 900  
 DB |||||  
 QY 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNERRLQKQVCNVSYPFF 900  
 DB |||||  
 QY 901 RAKAKVAFELDSFESKSIPLHLLEIELAAGSDNERDSTKEDNAPLRFHLKYADVLFT 960  
 DB |||||  
 QY 901 RAKAKVAFELDSFESKSIPLHLLEIELAAGSDNERDSTKEDNAPLRFHLKYADVLFT 960  
 DB |||||  
 QY 961 RSSLSHYEVKLNLSLERYDGIQPPFSCIFRIQNLGLPFIHGMMWKITPIATRSNRL 1020  
 DB |||||  
 QY 961 RSSLSHYEVKLNLSLERYDGIQPPFSCIFRIQNLGLPFIHGMMWKITPIATRSNRL 1020  
 DB |||||  
 QY 1021 KLDFLTDANTSCINWNSSTYRTPVVEEDLERAPOLNHSNDSVVSINCNIRLVPNQE 1080  
 DB |||||  
 QY 1021 KLDFLTDANTSCINWNSSTYRTPVVEEDLERAPOLNHSNDSVVSINCNIRLVPNQE 1080  
 DB |||||  
 QY 1081 NFHLGLNLWLSRLKALKYKSMIMVNAALQROFHPSPFIREDPSRQIEFEISKQEDQV 1140  
 DB |||||  
 QY 1081 NFHLGLNLWLSRLKALKYKSMIMVNAALQROFHPSPFIREDPSRQIEFEISKQEDQV 1140  
 DB |||||  
 QY 1141 PIWIIIVGSTLGLLALLVLAALKLGFRRSARRRPGLDTPPKVLE 1188  
 DB |||||  
 QY 1141 PIWIIIVGSTLGLLALLVLAALKLGFRRSARRRPGLDTPPKVLE 1188  
 DB |||||  
 RESULT 7  
 ADE09956  
 ID ADE09956 standard; protein; 1188 AA.  
 AC ADE09956;  
 DT 29-JAN-2004 (first entry)  
 XX Novel protein-related contig polypeptide sequence #544.  
 XX novel gene; novel protein; tissue marker; molecular weight marker;  
 KW chromosome marker; Genetic disorder; contig.  
 XX Unidentified.  
 XX WO2003054152-A2.  
 XX 03-JUL-2003.  
 XX 10-DEC-2002; 2002WO-US039555.  
 XX 10-DEC-2001; 2001US-0339739P.  
 PR 11-DEC-2001; 2001US-0339453P.  
 PR 14-MAR-2002; 2002US-0365091P.  
 PR 14-MAR-2002; 2002US-0365384P.  
 PR 12-APR-2002; 2002US-0372381P.  
 PR 12-APR-2002; 2002US-0372615P.  
 PR 22-APR-2002; 2002US-00128558.  
 PR 24-APR-2002; 2002US-0376045P.  
 XX (HYSB-) HYSEQ INC.  
 XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;  
 PI Chosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;  
 PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;  
 XX WPI; 2003-569235/53.  
 XX New polynucleotides, useful for expressing recombinant proteins for  
 PT analysis, characterization or therapeutic use, or as markers for tissues  
 PT in which the corresponding protein is preferentially expressed.  
 XX Disclosure; SEQ ID NO 3022; 1177pp; English.  
 PS The invention comprises the amino acid and coding sequences of novel  
 CC proteins. The DNA and protein sequences of the invention are useful as:  
 CC markers for tissues in which the corresponding protein is preferentially

CC expressed; as molecular weight markers on gels; as chromosome markers or  
 CC tags; to identify chromosomes or to map related gene positions; and to  
 CC compare with endogenous DNA sequences in patients to identify potential  
 CC genetic disorders. The present amino acid sequence was used in the  
 CC exemplification of the invention.  
 XX  
 SQ Sequence 1188 AA;  
 Query Match 99.5%; Score 6192; DB 7; Length 1188;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 1183; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MDLPRGLVAVAWLSLWPGTDTFNMDTRKPRVTPGSTAPFGYVQQHDISGNKWLVVGA 60  
 DB 1 MDLPRGLVAVAWLSLWPGTDTFNMDTRKPRVTPGSTAPFGYVQQHDISGNKWLVVGA 60  
 QY 61 PLENGYQKTDGVYKCPVIEHNCCTKLNLGRVTLNSVSRKDNMRGLGLSLATNPXNSFLA 120  
 DB 61 PLENGYQKTDGVYKCPVIEHNCCTKLNLGRVTLNSVSRKDNMRGLGLSLATNPXNSFLA 120  
 QY 121 CSPLWSEHCSSYTTGMCVRVNSNFRFSTKVPALQRCQTYMDIVIVLQGSNSIYPWE 180  
 DB 121 CSPLWSEHCSSYTTGMCVRVNSNFRFSTKVPALQRCQTYMDIVIVLQGSNSIYPWE 180  
 QY 181 VQFELINILKKFYIGPGQIQGVVQYGEDVVFHFLNDYRSVKDWVBAASHIEQGGTET 240  
 DB 181 VQFELINILKKFYIGPGQIQGVVQYGEDVVFHFLNDYRSVKDWVBAASHIEQGGTET 240  
 QY 241 RTAFGIEFARSEAPQKGRGAKKVMITVITDGHSDSPDLKVIQOESRDNVTRYAVVL 300  
 DB 241 RTAFGIEFARSEAPQKGRGAKKVMITVITDGHSDSPDLKVIQOESRDNVTRYAVVL 300  
 QY 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVITDEAALKDVIDALGDRIFSLCTNKNET 360  
 DB 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVITDEAALKDVIDALGDRIFSLCTNKNET 360  
 QY 361 SFGLEMSQTFSSHVEDGVLLGAVGAYDNGAVLKGTSAGKVIPLRESYLKEPPELKN 420  
 DB 361 SFGLEMSQTFSSHVEDGVLLGAVGAYDNGAVLKGTSAGKVIPLRESYLKEPPELKN 420  
 QY 421 HGAYLVTVTSVSSRQGRVYVAGAPFNHTGKVLFTMNNSRLTIHQAMRQOIGSYF 480  
 DB 421 HGAYLVTVTSVSSRQGRVYVAGAPFNHTGKVLFTMNNSRLTIHQAMRQOIGSYF 480  
 QY 481 GSEITSDIDGDGVTDLVLCAPMYFNEGRERGKVVYVELRQNRPFVNGTLKDSHSYQNA 540  
 DB 481 GSEITSDIDGDGVTDLVLCAPMYFNEGRERGKVVYVELRQNRPFVNGTLKDSHSYQNA 540  
 QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGRGSLTKTPKORITASELATG 600  
 DB 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGRGSLTKTPKORITASELATG 600  
 QY 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNAVILWSRPVQINASLHFEPSKINIFHRDCK 660  
 DB 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNAVILWSRPVQINASLHFEPSKINIFHRDCK 660  
 QY 661 RSGRDATCLAFCTPIELAPHFQTTVGIRYNATMDERRYPRAHLDEGGDFTRAV 720  
 DB 661 RSGRDATCLAFCTPIELAPHFQTTVGIRYNATMDERRYPRAHLDEGGDFTRAV 720  
 QY 721 LLSGQELCERINFHVLDTADYVKPVTFSVEYSLEDPDHGMPLDDGPTTLRVSPFWNG 780  
 DB 721 LLSGQELCERINFHVLDTADYVKPVTFSVEYSLEDPDHGMPLDDGPTTLRVSPFWNG 780  
 QY 781 CNEDEHCVPLVDARSDDLPTAMEYCORVLKPAQDCSAYTSLSDTTVFIIESTRQAV 840  
 DB 781 CNEDEHCVPLVDARSDDLPTAMEYCORVLKPAQDCSAYTSLSDTTVFIIESTRQAV 840  
 QY 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNERRLQKQVCNVSYPFF 900  
 DB 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNERRLQKQVCNVSYPFF 900  
 QY 901 RAKAKVAFRLDSFESKSIPLHLLEIELAAGSDNERDSTKEDNAPLRFHLKYADVLFT 960



Db 781 CNEDEHCVPDLVDARSDDLPTAMEYCORVLRPAQDCSAYTUSFDFTVFIISSTQRVAV 840  
Qy 841 EATLENGENAYSTVLNISOANLQFASLTQKEDSDGSIQECVNEERLQKQVNSYPPF 900  
Db 841 EATLENGENAYSTVLNISOANLQFASLTQKEDSDGSIQECVNEERLQKQVNSYPPF 900  
Qy 901 RAKAKVAFRLDSFSSKIFLHLEIELAAGSDSNERDSTKEDNVAFLRPHLYEADVLPT 960  
Db 901 RAKAKVAFRLDSFSSKIFLHLEIELAAGSDSNERDSTKEDNVAFLRPHLYEADVLPT 960  
Qy 961 RSSLSHYEVKLNSSLYRDVIGIPFPSCIFRIQNLGLFPIHGMKTIPIATRSNRL 1020  
Db 961 RSSLSHYEVKLNSSLYRDVIGIPFPSCIFRIQNLGLFPIHGMKTIPIATRSNRL 1020  
Qy 1021 KLRDPLTDE-ANTSCINWGNSTYRTPVEEDLRRAPQLNHSNDVSNIRLVPNOE 1079  
Db 1021 KLRDPLTDE-ANTSCINWGNSTYRTPVEEDLRRAPQLNHSNDVSNIRLVPNOE 1080  
Qy 1080 INFHLLGNLWRLSLKALKYKSMKIMVNAALQROFHSPPFIREDDPSRQIEFEISKQEDWQ 1139  
Db 1081 INFHLLGNLWRLSLKALKYKSMKIMVNAALQROFHSPPFIREDDPSRQIEFEISKQEDWQ 1140  
Qy 1140 VPIWIIIVGSTLGGILLIALLVLALRKLGFPRSRARRRREPGLDPTPKVLE 1188  
Db 1141 VPIWIIIVGSTLGGILLIALLVLALRKLGFPRSRARRRREPGLDPTPKVLE 1189

RESULT 9  
ABR58364 standard; protein; 1189 AA.  
ID ABR58364;  
XX AC ABR58364;  
XX DT 07-JUL-2003 (first entry)  
XX DE Human NOV2a.  
XX Kw Human; NOV; antidiabetic; anorectic; antibacterial; virucide;  
Kw immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia;  
Kw antiparkinsonian; antilipaeamic; gene therapy; metabolic disorder;  
Kw diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;  
Kw neurodegenerative disorder; Alzheimer's disease; immune disorder;  
Kw haematopoietic disorder.  
XX OS Homo sapiens.  
XX PN WO2003029423-A2.  
XX PD 10-APR-2003.  
XX PF 02-OCT-2002; 2002WO-US031358.  
XX 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327342P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339266P.  
PR 29-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346337P.  
PR 12-APR-2002; 2002US-0371972P.  
PR 12-APR-2002; 2002US-0371980P.  
PR 17-APR-2002; 2002US-0373261P.  
PR 19-APR-2002; 2002US-0373805P.  
PR 23-APR-2002; 2002US-0374738P.  
PR 16-MAY-2002; 2002US-0381101P.

PR 17-MAY-2002; 2002US-0381635P.  
PR 29-MAY-2002; 2002US-0383830P.  
PR 01-OCT-2002; 2002US-00262839.  
XX (CURA-) CURAGEN CORP.  
XX Alsobrook JP, Ellerman DW, Boldog PL, Burgess CE, Catterton E;  
PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;  
PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;  
PI Rothenberg ME, Shinkens RA, Smithson G, Spytek KA, Taupier RJ;  
PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;  
XX WPI; 2003-381625/36.  
DR N-PSDB; ACC72076.  
XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or  
PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or  
PT dyslipidemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX Claim 1; Page 105; 487pp; English.  
XX The present invention relates to novel human NOV proteins and their  
CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV  
CC proteins are useful in manufacturing a medicament for treating a syndrome  
CC associated with a human disease. The NOV proteins and coding sequences  
CC may be used to diagnose, treat or prevent metabolic disorders such as  
CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative  
CC disorders such as Alzheimer's disease or Parkinson's disease, immune  
CC disorders, haematopoietic disorders and various dyslipidaemias  
XX Sequence 1189 AA;  
SQ Query Match 99.4%; Score 6188.5; DB 6; Length 1189;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1184; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
Qy 1 MDLPGLVAVNALSILWPGFTDTFNMOTKPRVIPCGRRTAFGTYTQOHDISGNKVLVGA 60  
Db 1 MDLPGLVAVNALSILWPGFTDTFNMOTKPRVIPCGRRTAFGTYTQOHDISGNKVLVGA 60  
Qy 61 PLEITNGYQKTVGVYKCPVHGNCTKLNLRVTLNSVSRKONMRLGLSLATPKDNSFLA 120  
Db 61 PLEITNGYQKTVGVYKCPVHGNCTKLNLRVTLNSVSRKONMRLGLSLATPKDNSFLA 120  
Qy 121 CSPLMSHECGSSYYTTCMSRVNSNFRSKTVAPALQRCQTYMDIVIVLDSNSIYPWVE 180  
Db 121 CSPLMSHECGSSYYTTCMSRVNSNFRSKTVAPALQRCQTYMDIVIVLDSNSIYPWVE 180  
Qy 181 VOHFLINILKPYIGPGQIQGVQYGEDVHVEFHNDYRSVKDVVEAASHIEQGGTET 240  
Db 181 VOHFLINILKPYIGPGQIQGVQYGEDVHVEFHNDYRSVKDVVEAASHIEQGGTET 240  
Qy 241 RTAPGIEFARSAFQKGRKGAKKVMIVITDGEHSDSPDLKVIQOOSRDNTRYAVAVL 300  
Db 241 RTAPGIEFARSAFQKGRKGAKKVMIVITDGEHSDSPDLKVIQOOSRDNTRYAVAVL 300  
Qy 301 GYNNRGINPETFLNEIKYIASDDPKHFNVTDEAALKDIYDALGDRIFSLEGTKNET 360  
Db 301 GYNNRGINPETFLNEIKYIASDDPKHFNVTDEAALKDIYDALGDRIFSLEGTKNET 360  
Qy 361 SFGLEMSQTGFSHVVEDGVLGAYDNGAVLKETSAGKVIPLRSYLKEFPPELKN 420  
Db 361 SFGLEMSQTGFSHVVEDGVLGAYDNGAVLKETSAGKVIPLRSYLKEFPPELKN 420  
Qy 421 HGAYLGYTVTSVSSRQGEVYVAGAPRNFHTGKVLFTMHNRSLSLIHQAMGQQIGSYF 480  
Db 421 HGAYLGYTVTSVSSRQGEVYVAGAPRNFHTGKVLFTMHNRSLSLIHQAMGQQIGSYF 480  
Qy 481 GSEITTSVIDIDGQVTDVLLVAGAPMYFNEGRGRGVYVYELRQNRFYNTGTLKDSHSYQNA 540  
Db 481 GSEITTSVIDIDGQVTDVLLVAGAPMYFNEGRGRGVYVYELRQNRFYNTGTLKDSHSYQNA 540

|    |      |  |            |           |                             |          |        |      |
|----|------|--|------------|-----------|-----------------------------|----------|--------|------|
| Qy | 541  | RFSSIASVRDLNODSYNDVVVGHAPLEDNHAGAIYI   | FHGFRGSI   | LKTPKQIR  | TASHIATG                    | 600      |        |      |
| Db | 541  | RFSSIASVRDLNODSYNDVVVGHAPLEDNHAGAIYI   | FHGFRGSI   | SLKTPKQIR | TASEIATG                    | 600      |        |      |
| Qy | 601  | LQYFGCSIHGOLDLNEGLDLAVAGALGNAILMSRPVQV | INASLHSE   | PEPSKINI  | FHRDCK                      | 660      |        |      |
| Db | 601  | LQYFGCSIHGOLDLNEGLDLAVAGALGNAILMSRPVQV | INASLHSE   | PEPSKINI  | FHRDCK                      | 660      |        |      |
| Qy | 661  | RSGRDACLAAFLCFTPIFLAPHQFTTVTGIRYNAT    | MDERRYT    | TPRAHL    | DGCGDRFTNRAV                | 720      |        |      |
| Db | 661  | RSGRDACLAAFLCFTPIFLAPHQFTTVTGIRYNAT    | MDERRYT    | TPRAHL    | DGCGDRFTNRAV                | 720      |        |      |
| Qy | 721  | LLSGQBCLERINPHVLDTADYVVKPFTFSVEYS      | LEDPDHGP   | MLDDG     | WPTTLIRVSVPPFWG             | 780      |        |      |
| Db | 721  | LLSGQBCLERINPHVLDTADYVVKPFTFSVEYS      | LEDPDHGP   | MLDDG     | WPTTLIRVSVPPFWG             | 780      |        |      |
| Qy | 781  | CNEDEHCVPDVLVDARSDLPTAMEYCORVLKPAQDC   | SAYT       | TLSPDT    | TVFIIESTRQRVAV              | 840      |        |      |
| Db | 781  | CNEDEHCVPDVLVDARSDLPTAMEYCORVLKPAQDC   | SAYT       | TLSPDT    | TVFIIESTRQRVAV              | 840      |        |      |
| Qy | 841  | EATLENGENAYSTVLINISQSANLOFASL          | LOKEDSDGSI | ECVNER    | BLQKQVCNVSYPFF              | 900      |        |      |
| Db | 841  | EATLENGENAYSTVLINISQSANLOFASL          | LOKEDSDGSI | ECVNER    | BLQKQVCNVSYPFF              | 900      |        |      |
| Qy | 901  | RAKAKVAFRLDSEPSKSI                     | FIHLHLEI   | LAAGSD    | SNERDSTKEDNVAPLRFHLKYEADVLT | 960      |        |      |
| Db | 901  | RAKAKVAFRLDSEPSKSI                     | FIHLHLEI   | LAAGSD    | SNERDSTKEDNVAPLRFHLKYEADVLT | 960      |        |      |
| Qy | 961  | RSSLSHYEVKLSNLSERYDIGPFPFCI            | FRIQNLG    | LFPFHGM   | MMKTIIFIA                   | TRSGNRLL | 1020   |      |
| Db | 961  | RSSLSHYEVKLSNLSERYDIGPFPFCI            | FRIQNLG    | LFPFHGM   | MMKTIIFIA                   | TRSGNRLL | 1020   |      |
| Qy | 1021 | KLDRFLHDE-ANTSCNIWGNSTERYPTFVEED       | LRRAPQL    | NHNSDDVYS | INCNI                       | RLVPNOE  | 1079   |      |
| Db | 1021 | KLDRFLHDE-ANTSCNIWGNSTERYPTFVEED       | LRRAPQL    | NHNSDDVYS | INCNI                       | RLVPNOE  | 1080   |      |
| Qy | 1080 | INPHLLGNLWLSRKALKYKSMKIMVNAALQ         | QPHSPFI    | FREEDP    | SPQI                        | SEFIS    | KQSDMQ | 1139 |
| Db | 1081 | INPHLLGNLWLSRKALKYKSMKIMVNAALQ         | QPHSPFI    | FREEDP    | SPQI                        | VFEIS    | KQSDMQ | 1140 |
| Qy | 1140 | VPIWIIVGSTLGGLLALLVIALKXLF             | GFPSARR    | REPGLD    | TPPKVLE                     | 1188     |        |      |
| Db | 1141 | VPIWIIVGSTLGGLLALLVIALKXLF             | GFPSARR    | REPGLD    | TPPKVLE                     | 1189     |        |      |

## BEST COPY AVAILABLE

RESULI IO  
ADA27054  
ID ADA27054 standard: protein: 1189 AA.

AC ADA27054:

20-NOV-2003 (first entry)

XX DE Human novel secreted protein from cDNA HOHBV69 #1.

cytostatic; antiinflammatory; immunomodulator; neuroprotective;  
 hemostatic; gene therapy; cancer; inflammation; immune disorder;  
 neurological disorder; blood clotting disorder; food additive;  
 preservative; human; secreted protein.

Homo sapiens.

XX PN US2003055231-A1.

XX  
20-MAR-2003

29-QCT-2001: 2001US-00984130

XX  
BB 20-007-1000. 0078-01050710

PR 27-OCT-1999; 99WO-US025031.

PR 30-OCT-2000; 2000US-0243792P.

NY 92-1077, 100-801007, 100-801007

XX (NIJ/J/ NI J.  
PA {VOUN/} YOUNG P E.  
PA {KENN/} KENNY J J.  
PA {OLSE/} OLSEN H S.  
PA {MOOR/} MOORE P A.  
PA {WEIV/} WEI Y.  
PA {GREB/} GREENE J M.  
PA {RUBE/} RUBEN S M.  
PA {LIUD/} LIU D.  
PA {CROC/} CROCKER P R.

PI Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;  
PI Ruben SM, Liu D, Crocker PR;

DR WPI; 2003-567103/53.  
DR N-PSDB; ADA27036.

AA New human secreted nucleic acid molecules and polypeptides, useful for  
PT preventing, treating, or ameliorating a medical condition, such as  
PT cancer, inflammation, immune disorders, neurological and blood clotting  
PT disorders.

xx  
ps  
Claim 11: Fig 19: 454pp: English:

The invention relates to an isolated nucleic molecule that is at least 95% identical to 18 human cDNA sequences representing 12 novel genes encoding secreted proteins or a polynucleotide fragment of the cDNA sequence contained in American Type Culture Collection (ATCC) deposit No. defined in the specification, its species homologue, a variant or allelic variant of the polynucleotide having a polynucleotide capable of hybridising under conditions the polynucleotide, where the polynucleotide does not hybridise under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A or T residues. Also included are recombinant vectors, host cells for producing the polypeptide, the secreted polypeptide (comprising a sequence that is at least 95% identical to a polypeptide fragment, domain, epitope, full-length protein, variant, allelic variant or species homologue), antibodies that specifically bind to the polypeptides, diagnosing, treating, preventing or ameliorating a medical condition by administering the polynucleotide or the polypeptide, the gene corresponding to the cDNA sequence and identifying an activity in a biological assay (by expressing the cDNA sequence in a cell, isolating the supernatant, and detecting an activity in a biological assay and identifying the protein in the supernatant having the activity). The polypeptides, nucleic acids and antibodies are useful for diagnosing a pathological condition or a susceptibility to a pathological condition, for preventing, treating, or ameliorating a medical condition, such as cancer, inflammation and other immune disorders, neurological and blood clotting disorders (many examples are given in the specification). The nucleic acids are also useful for chromosome identification, radiation hybrid mapping or long-range restriction mapping. The polypeptides and antibodies are useful for providing immunological probes for differential identification of the tissues immunohistochemistry assays. The polypeptide, polynucleotide, agonist or antagonist may also be used as a food additive or preservative to increase or decrease storage capabilities, fat content or other nutritional components. The present is a secreted protein of the invention.

Sequence 1189 AA;

|                       |              |              |            |             |
|-----------------------|--------------|--------------|------------|-------------|
| Query Match           | 99.4%        | Score 6188.5 | DB 6       | Length 1189 |
| Best Local Similarity | 99.6%        | Pred. No. 0  |            |             |
| Matches 1184:         | Conservative | 1            | Mismatches | 3           |
|                       |              |              | Indels     | 1           |
|                       |              |              | Gaps       | 1           |

1 MDLPRLVVAWALSILWPGETDTENMDTRKPRVPGSRTAFEGYTVOOHDISGNKWL VVGA 60

db 1 MTI.PRG.LV.AW.ALS.LWP.GE.TD.TEN.MD.TR.KP.RV.I.PG.SR.TA.FG.YT.VO.OH.DI.SG.NK.WL.VV.GA.60

61 BIPPTNGVOKTCTNYKCRIFHCNCTKLNIGPVITI.SNYSERKDNMRI GLSIA TNPKNSELA 120

61 PI ETNYOVKTGNYKCRPHHCHNCTKI NI GRVTL SNVSEKDNMRI GLSLATNPKNSELA 120



Qy 121 CSPWLSHECGSSYTTTCGSRVNSNPRFSKTVAPALQRCQCTYMDIVIVLDGNSIYPWVE 180  
Db |||||  
Qy 121 CSPWLSHECGSSYTTTCGSRVNSNPRFSKTVAPALQRCQCTYMDIVIVLDGNSIYPWVE 180  
Db |||||  
Qy 181 VQHFLNLIKKEFIIGPQOIQGVVQVGEDVHVEFHNDYRSKDVVEAAASHIEQRGTET 240  
Db |||||  
Qy 181 VQHFLNLIKKEFIIGPQOIQGVVQVGEDVHVEFHNDYRSKDVVEAAASHIEQRGTET 240  
Db |||||  
Qy 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSHSDSPDLKVIQSERDNTVRYAVVL 300  
Db |||||  
Qy 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSHSDSPDLKVIQSERDNTVRYAVVL 300  
Db |||||  
Qy 301 GYNRGINPETFLNEIKYTASDPDKHFNVTDEAALKDIDVALGDRIPSEGTKNET 360  
Db |||||  
Qy 301 GYNRGINPETFLNEIKYTASDPDKHFNVTDEAALKDIDVALGDRIPSEGTKNET 360  
Db |||||  
Qy 361 SFGLEMSQTSFHHVVDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPELKN 420  
Db |||||  
Qy 361 SFGLEMSQTSFHHVVDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPELKN 420  
Db |||||  
Qy 421 HGAYLGYTTSVSVSSRQRYVYVAGAPFNHTGKVLPTMNNRSLTIHQAMRQOQIGSYF 480  
Db |||||  
Qy 421 HGAYLGYTTSVSVSSRQRYVYVAGAPFNHTGKVLPTMNNRSLTIHQAMRQOQIGSYF 480  
Db |||||  
Qy 481 GSEITSVDIDGQVTDVLLVGAEMYFNEGRERKGVVYVELRQNFVYNGTLKDSHSYQNA 540  
Db |||||  
Qy 481 GSEITSVDIDGQVTDVLLVGAEMYFNEGRERKGVVYVELRQNFVYNGTLKDSHSYQNA 540  
Db |||||  
Qy 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNEAGAIYIFHGRGSIILKTPKQITASELATG 600  
Db |||||  
Qy 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNEAGAIYIFHGRGSIILKTPKQITASELATG 600  
Db |||||  
Qy 601 LQVFGCSIHGOLDNLDELAVGALGNVILMSRPVQINASLHFEPSKINIFHRDCK 660  
Db |||||  
Qy 601 LQVFGCSIHGOLDNLDELAVGALGNVILMSRPVQINASLHFEPSKINIFHRDCK 660  
Db |||||  
Qy 661 RSGEDATCLAAFLCFTPIFLAPHFQITTVGIRYNATMDERRYTPRAHLDGGDRFTNRAV 720  
Db |||||  
Qy 661 RSGEDATCLAAFLCFTPIFLAPHFQITTVGIRYNATMDERRYTPRAHLDGGDRFTNRAV 720  
Db |||||  
Qy 721 LLSGQBLERINHFHLDTDADYKPTFSVEYSLDHPDHPGMDGQNPITLRSVFPFWNG 780  
Db |||||  
Qy 721 LLSGQBLERINHFHLDTDADYKPTFSVEYSLDHPDHPGMDGQNPITLRSVFPFWNG 780  
Db |||||  
Qy 781 CNDDEHCVPLDLARSDLPTAMEYCORVLKPAQDCSAYTLGFDFTVFIESTRQVAV 840  
Db |||||  
Qy 781 CNDDEHCVPLDLARSDLPTAMEYCORVLKPAQDCSAYTLGFDFTVFIESTRQVAV 840  
Db |||||  
Qy 841 EATLENGENAYSTVLNISOSANLQFASLIQKEDSDGSIIECVNEERLQKQVNSYPPF 900  
Db |||||  
Qy 841 EATLENGENAYSTVLNISOSANLQFASLIQKEDSDGSIIECVNEERLQKQVNSYPPF 900  
Db |||||  
Qy 901 RAKAKAFRLDSDRFSKIFLHLEIELAAGSDSNRSDTKEDNVAPLRPHLKYEAADVLT 960  
Db |||||  
Qy 901 RAKAKAFRLDSDRFSKIFLHLEIELAAGSDSNRSDTKEDNVAPLRPHLKYEAADVLT 960  
Db |||||  
Qy 961 RSSLSHYEKLNSLERYDGIQPPFSCIPIQNLGLFPIHGMKMTIPIATRSNRL 1020  
Db |||||  
Qy 961 RSSLSHYEKLNSLERYDGIQPPFSCIPIQNLGLFPIHGMKMTIPIATRSNRL 1020  
Db |||||  
Qy 1021 KLDPFLTDE-ANTSCTNIGWSTYRPTPVEEDLRAPOLNHSNDVVSINCNRLVPNOE 1079  
Db |||||  
Qy 1021 KLDPFLTDE-ANTSCTNIGWSTYRPTPVEEDLRAPOLNHSNDVVSINCNRLVPNOE 1080  
Db |||||  
Qy 1080 INFHLGNLWLRSLKALKYKSMKIMVNAALQROFHSPPFIFREEDPSRQIEFESKQEDWQ 1139  
Db |||||  
Qy 1081 INFHLGNLWLRSLKALKYKSMKIMVNAALQROFHSPPFIFREEDPSRQIEFESKQEDWQ 1140  
Db |||||  
Qy 1140 VPIWIIVGSITGLGILLIALLVLRKLGFPSRRRRREPGLDPTPKVLE 1188  
Db |||||  
Qy 1141 VPIWIIVGSITGLGILLIALLVLRKLGFPSRRRRREPGLDPTPKVLE 1189  
Db |||||

## RESULT 11

ADE63570

ID ADE63570 standard; protein; 1189 AA.

XX ADE63570;

XX DT 29-JAN-2004 (first entry)

XX DE Human Protein Q9UKX5, SEQ ID NO 9514.

XX KW Human; pain; neuronal tissue; gene therapy;

XX KW spinal segmental nerve injury; chronic constriction injury; CCI;

XX KW spared nerve injury; SNI; Chung.

XX OS Homo sapiens.

PN WO2003016475-A2.

PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-033347P.

XX (GHEO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX DR WPI; 2003-268312/26.

XX PT New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1189 AA;

Query Match 99.4%; Score 6188.5; DB 7; Length 1189;  
Best Local Similarity 99.6%; Fred. No. 0;  
Matches 1184; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 1 MDLPRGLVAVWALSINLPGQTDFTFNMDTRKPRVPGSRPTAFYTYVQOHDHSGNKLWVGA 60

|    |      |  |      |
|----|------|--|------|
| Db | 1    | MDLPRGVLVVAWLSLWPGFTDFNNDRTRKPRVIFGSRFAFGYTVQOEDTSGNKMVLVGA    | 60   |
| Qy | 61   | PLETNGYQKTGDVYKCPVIHGNCTKJNLGRVTLNWSERKNMELGLSLATNPKDNSFLA     | 120  |
| Db | 61   | PLETNGYQKTGDVYKCPVIHGNCTKJNLGRVTLNWSERKNMELGLSLATNPKDNSFLA     | 120  |
| Qy | 121  | CSPJWSHECGSSYYTTGMCNRVNSNFPFSKTVAPALORCQTYXDIIVIVLPGSNSIYFWZE  | 180  |
| Db | 121  | CSPJWSHECGSSYYTTGMCNRVNSNFPFSKTVAPALORCQTYXDIIVIVLPGSNSIYFWZE  | 180  |
| Qy | 181  | VQHFPLINILKKFYIGPGQIOGVVQYQGBDVVHBFHLNDYRSVKDVAASHIEORGTTET    | 240  |
| Db | 181  | VQHFPLINILKKFYIGPGQIOGVVQYQGBDVVHBFHLNDYRSVKDVAASHIEORGTTET    | 240  |
| Qy | 241  | RTAFGIEFARSEAFQKGRKGAKKVMIVITDGESHSDPDLEKVIQQSERDNVTYAYAVL     | 300  |
| Db | 241  | RTAFGIEFARSEAFQKGRKGAKKVMIVITDGESHSDPDLEKVIQQSERDNVTYAYAVL     | 300  |
| Qy | 301  | GYNRRGINPBTFLNEIKYJASDPDDKHFFNVTVDEALKDIDVALGDRIIPSLGNTKNKET   | 360  |
| Db | 301  | GYNRRGINPBTFLNEIKYJASDPDDKHFFNVTVDEALKDIDVALGDRIIPSLGNTKNKET   | 360  |
| Qy | 361  | SFGLMSQTFSSHVJEDGVLLGAVGAYDMNGAVLKETSAKVILPRLSYLKBFPBELKN      | 420  |
| Db | 361  | SFGLMSQTFSSHVJEDGVLLGAVGAYDMNGAVLKETSAKVILPRLSYLKBFPBELKN      | 420  |
| Qy | 421  | HGAYLGYTVTVSSSROGRVYVAGAPFNHTGKVILFTMHNRSILTHQWRGQOIGSYF       | 480  |
| Db | 421  | HGAYLGYTVTVSSSROGRVYVAGAPFNHTGKVILFTMHNRSILTHQWRGQOIGSYF       | 480  |
| Qy | 481  | GSEITTSVDIDGQVTDVLLNGAPMYFNEGREGKVVYVELQNEFVNGTLKDSHSYQNA      | 540  |
| Db | 481  | GSEITTSVDIDGQVTDVLLNGAPMYFNEGREGKVVYVELQNEFVNGTLKDSHSYQNA      | 540  |
| Qy | 541  | RFGSSIASVRDLNQSNDVVVGAPLBDNHAGALYIFHGFGRSILTKPKOSITASELATG     | 600  |
| Db | 541  | RFGSSIASVRDLNQSNDVVVGAPLBDNHAGALYIFHGFGRSILTKPKOSITASELATG     | 600  |
| Qy | 601  | LOYFGCSIHGOLDINEDGLIDLAVGALGNVILNSRPVQVQINASHIHPSPKINIFHRDCK   | 660  |
| Db | 601  | LOYFGCSIHGOLDINEDGLIDLAVGALGNVILNSRPVQVQINASHIHPSPKINIFHRDCK   | 660  |
| Qy | 661  | RSGRDATCLAAFLCFTPIFLAPHFQTTVIGIRYNATMDERRYTPRAHLDGEGDRFTNRV    | 720  |
| Db | 661  | RSGRDATCLAAFLCFTPIFLAPHFQTTVIGIRYNATMDERRYTPRAHLDGEGDRFTNRV    | 720  |
| Qy | 721  | LLSSQBLCEINHFHVLDTADYKPVTFSEVYSLEDDPHGPMIDGHPPTILRVSPFWNG      | 780  |
| Db | 721  | LLSSQBLCEINHFHVLDTADYKPVTFSEVYSLEDDPHGPMIDGHPPTILRVSPFWNG      | 780  |
| Qy | 781  | CNEDEHCVPDLVDLARSDLPTAMEYQCORVLKPKAQDCSAYTSLSEDFTTVFIIESTRQVAV | 840  |
| Db | 781  | CNEDEHCVPDLVDLARSDLPTAMEYQCORVLKPKAQDCSAYTSLSEDFTTVFIIESTRQVAV | 840  |
| Qy | 841  | EATLENGENAYSTVLNISQSANLOFASLIQKEDSDGSIECVNEERLQKQVCNVSYPFF     | 900  |
| Db | 841  | EATLENGENAYSTVLNISQSANLOFASLIQKEDSDGSIECVNEERLQKQVCNVSYPFF     | 900  |
| Qy | 901  | RAKAKVAPRLDSEFSKSIFFLHLHILSLAAGSDSNERDSTKEDNVAPLPHLKYEADVLFT   | 960  |
| Db | 901  | RAKAKVAPRLDSEFSKSIFFLHLHILSLAAGSDSNERDSTKEDNVAPLPHLKYEADVLFT   | 960  |
| Qy | 961  | RSSLSLHYEVKLNSESLERYDGGIPPPFSCIPRIQNLGJFFPIHGMMWKITIPATRSQNRLL | 1020 |
| Db | 961  | RSSLSLHYEVKLNSESLERYDGGIPPPFSCIPRIQNLGJFFPIHGMMWKITIPATRSQNRLL | 1020 |
| Qy | 1021 | KLRDPLTDEANTSCNTWGNSTBYRPTPVEEDLRAAPOLNHSNDVVSINCLRLVPNOE      | 1079 |
| Db | 1021 | KLRDPLTDEVANTSCNTWGNSTBYRPTPVEEDLRAAPOLNHSNDVVSINCLRLVPNOE     | 1080 |
| Qy | 1080 | INFLLGNLMRLSKALKYKSMKIMVNAALQROFSPFPIFREEDPSQIEFISQBDWQ        | 1139 |

[illegible]



Db 61 PLETNGYQKTDVYKCPVHGNCTKLNGLRVTLSNVSEKDNMRLGLSLATNPKDNSFLA 120  
Qy 121 CSPLWSHECCSSYTTGMCGRVNSNFRFSKTVAPALQRCOTYMDIVIVLDGNSIYPWVE 180  
Db 121 CSPLWSHECCSSYTTGMCGRVNSNFRFSKTVAPALQRCOTYMDIVIVLDGNSIYPWVE 180  
Qy 181 VOHFLNLIKFXIYIPQIQVGVVQYGEDVHVFHFLNDYVSVDVVEAASHIEQRCGTET 240  
Db 181 VOHFLNLIKFXIYIPQIQVGVVQYGEDVHVFHFLNDYVSVDVVEAASHIEQRCGTET 240  
Qy 241 RTAFGEFARSEAFQKGRGAKKVMIVITDGHSDSPDLEKVIQOSERDNVTRYAVAVL 300  
Db 241 RTAFGEFARSEAFQKGRGAKKVMIVITDGHSDSPDLEKVIQOSERDNVTRYAVAVL 300  
Qy 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVDEAALKDIDVALGDRIFSLSGTNKNET 360  
Db 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVDEAALKDIDVALGDRIFSLSGTNKNET 360  
Qy 361 SPGLEMSQTGFSSHVVDGVLGNVGYDNGAVLAKETSAGKVIPIRESVLAKEFPBELKN 420  
Db 361 SPGLEMSQTGFSSHVVDGVLGNVGYDNGAVLAKETSAGKVIPIRESVLAKEFPBELKN 420  
Qy 421 HGAYLGYTVTVSVSSRCGRVYVAGAPRFNHTGKVIPTMNNRSLTIHQAMRQQIGSYF 480  
Db 421 HGAYLGYTVTVSVSSRCGRVYVAGAPRFNHTGKVIPTMNNRSLTIHQAMRQQIGSYF 480  
Qy 481 GSEITSDIDGCGTVDVLLVCAQPMYFNEGRERKQVYVYELRQNFYNGTLKDSHXYQNA 540  
Db 481 GSEITSDIDGCGTVDVLLVCAQPMYFNEGRERKQVYVYELRQNFYNGTLKDSHXYQNA 540  
Qy 541 RFGSSIASVRDLNQDSYNDVVVGAFLDNGHAGAIYIFHGFGRGILKTPKORITASELATG 600  
Db 541 RFGSSIASVRDLNQDSYNDVVVGAFLDNGHAGAIYIFHGFGRGILKTPKORITASELATG 600  
Qy 601 LQYFGCSIHQCLDNLNEDGLDLAVGALGNVAVILWSPRVVQINASLHFEPSKINIFHRDCK 660  
Db 601 LQYFGCSIHQCLDNLNEDGLDLAVGALGNVAVILWSPRVVQINASLHFEPSKINIFHRDCK 660  
Qy 661 RSGRDAZCLAFLCFPIFLAPHFQTTVGIRVYATWDERRYTPRAHLDGGRDFNRVAV 720  
Db 661 RSGRDAZCLAFLCFPIFLAPHFQTTVGIRVYATWDERRYTPRAHLDGGRDFNRVAV 720  
Qy 721 LLSGGQELCBRIHFVLDADSLTADYVYKVPVTSVEYSLEDPDHGFMDDGMPTLRVSVFPWNG 780  
Db 721 LLSGGQELCBRIHFVLDADSLTADYVYKVPVTSVEYSLEDPDHGFMDDGMPTLRVSVFPWNG 780  
Qy 781 CNEDEHCVPLVDLADSLTADYVYKVPVTSVEYSLEDPDHGFMDDGMPTLRVSVFPWNG 840  
Db 781 CNEDEHCVPLVDLADSLTADYVYKVPVTSVEYSLEDPDHGFMDDGMPTLRVSVFPWNG 840  
Qy 841 EATLERNGENAYSTVLNISQSANLQFASLIQKEDSDGSEICVNEERLQKQVNCVSYPPF 900  
Db 841 EATLERNGENAYSTVLNISQSANLQFASLIQKEDSDGSEICVNEERLQKQVNCVSYPPF 900  
Qy 901 RAKAKVAFRLDSEFSSKIFLHLEIELAGSDNSNERDSTKEDNVAPLPHLYKYEADVLET 960  
Db 901 RAKAKVAFRLDSEFSSKIFLHLEIELAGSDNSNERDSTKEDNVAPLPHLYKYEADVLET 960  
Qy 961 RSSLSHYEVLKSSSLERYDGIQPPFSCIPRIQNLGFIPIHGMKMTIPIATRSQNRLL 1020  
Db 961 RSSLSHYEVLKSSSLERYDGIQPPFSCIPRIQNLGFIPIHGMKMTIPIATRSQNRLL 1020  
Qy 1021 KLKDFLTDE-ANTSCINWNGSTYRTPPVEEDLRRAPOLNHSNDSVVSNCNIRLVPNOE 1079  
Db 1021 KLKDFLTDE-ANTSCINWNGSTYRTPPVEEDLRRAPOLNHSNDSVVSNCNIRLVPNOE 1079  
Qy 1080 INFHLLGNLWLSLKALKYKSMIMVNAALORQFHSPFIPREDDPSQIREFISKEQDWQ 1139  
Db 1080 INFHLLGNLWLSLKALKYKSMIMVNAALORQFHSPFIPREDDPSQIREFISKEQDWQ 1139  
Qy 1140 VPIWIIWVSTGLGILLALLVLALRKLGFRRSARRRRPGLDPTPKVLE 1188  
Db 1141 VPIWIIWVSTGLGILLALLVLALRKLGFRRSARRRRPGLDPTPKVLE 1189

RESULT 13

ABR58365

ID ABR58365 standard; protein; 1120 AA.

XX

AC ABR58365;

XX

DT 07-JUL-2003 (first entry)

XX

DE Human NOV2b.

XX

KW Human; NOV; antidiabetic; anorectic; antibacterial; virucide;  
KW immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia;  
KW antiparkinsonian; antilipaeamic; gene therapy; metabolic disorder;  
KW diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;  
KW neurodegenerative disorder; Alzheimer's disease; immune disorder;  
KW haematopoietic disorder.

XX

OS Homo sapiens.

XX

PN WO2003029423-A2.

XX

PD 10-APR-2003.

XX

PF 02-OCT-2002; 2002WO-US031358.

XX

PR 02-OCT-2001; 2001US-0326483P.

PR

PR 05-OCT-2001; 2001US-0327342P.

PR

PR 09-OCT-2001; 2001US-0327917P.

PR

PR 09-OCT-2001; 2001US-0328029P.

PR

PR 09-OCT-2001; 2001US-0328044P.

PR

PR 09-OCT-2001; 2001US-0328056P.

PR

PR 12-OCT-2001; 2001US-0328849P.

PR

PR 15-OCT-2001; 2001US-0329414P.

PR

PR 17-OCT-2001; 2001US-0330142P.

PR

PR 22-OCT-2001; 2001US-0341058P.

PR

PR 24-OCT-2001; 2001US-0339266P.

PR

PR 24-OCT-2001; 2001US-0343629P.

PR

PR 29-OCT-2001; 2001US-0349575P.

PR

PR 01-NOV-2001; 2001US-0346357P.

PR

PR 12-APR-2002; 2002US-0371972P.

PR

PR 12-APR-2002; 2002US-0371980P.

PR

PR 17-APR-2002; 2002US-0373261P.

PR

PR 19-APR-2002; 2002US-0373805P.

PR

PR 23-APR-2002; 2002US-0374738P.

PR

PR 16-MAY-2002; 2002US-0381101P.

PR

PR 17-MAY-2002; 2002US-0381635P.

PR

PR 29-MAY-2002; 2002US-0383630P.

PR

PR 01-OCT-2002; 2002US-0026283P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;  
PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;  
PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;  
PI Rothenberg ME, Shinkets RA, Smithson G, Spytek KA, Taupier RJ;  
PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;

XX WPI; 2003-381625/36.

DR

DR N-PSDB; ACC72077.

XX

XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or  
XX treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or  
XX dyslipidemia, and in chromosome mapping, tissue typing or  
XX pharmacogenomics.

PT

PT Claim 1; Page 107; 487pp; English.

XX

XX The present invention relates to novel human NOV proteins and their  
XX coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV  
XX proteins are useful in manufacturing a medicament for treating a syndrome  
XX associated with a human disease. The NOV proteins and coding sequences

CC may be used to diagnose, treat or prevent metabolic disorders such as  
CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative  
CC disorders such as Alzheimer's disease or Parkinson's disease, immune  
CC disorders, haematopoietic disorders and various dyslipidaemias  
XX  
SQ Sequence 1120 AA;

Query Match 92.9%; Score 5780; DB 6; Length 1120;  
Best Local Similarity 93.8%; Pred. NO. 0;  
Matches 1115; Conservative 1; Mismatches 3; Indels 70; Gaps 2;

QY 1 MDLPRGLVAVALSLLWPGFTDTFNDTRKPRVIGSRRTAFPGYTVQQHDSGNKWLAVGA 60  
DB 1 MDLPRGLVAVALSLLWPGFTDTFNDTRKPRVIGSRRTAFPGYTVQQHDSGNKWLAVGA 60  
QY 61 PLENGYQKTDGVYKCPVIHGNCTKLNIGRVTLNVSVSRKDNMRGLSLATNPKDMSFLA 120  
DB 61 PLENGYQKTDGVYKCPVIHGNCTKLNIG----- 89  
QY 121 CSPLWSHSCGSYYTTCGRVNSFRSKTVAPALQRCOTYMDIVIVLDGNSNITYPWE 180  
DB 90 -----COTYMDIVIVLDGNSNITYPWE 111  
QY 181 VQHFLINILKKFYIGFGQIQGVVQYGEDVWHEHFLNDYRSVKCVVEAASHIEQGGTET 240  
DB 112 VQHFLINILKKFYIGFGQIQGVVQYGEDVWHEHFLNDYRSVKCVVEAASHIEQGGTET 171  
QY 241 RTAFGIEFARBEAFQKGRKAKKXVIVITGESHDSPLLEKVIQOQSRDNVTRYAVAVL 300  
DB 172 RTAFGIEFARBEAFQKGRKAKKXVIVITGESHDSPLLEKVIQOQSRDNVTRYAVAVL 231  
QY 301 GYNNRGINPETFNEIKYIASDPDKHFFNVYDEAALKDIDVALGDRIFSLGNTKNET 360  
DB 232 GYNNRGINPETFNEIKYIASDPDKHFFNVYDEAALKDIDVALGDRIFSLGNTKNET 291  
QY 361 SFGLEMSQTGSSSHVVEGVLLGAVGAYDNGAVLKETSAKVIPLRSEYKKEPPELUN 420  
DB 292 SFGLEMSQTGSSSHVVEGVLLGAVGAYDNGAVLKETSAKVIPLRSEYKKEPPELUN 351  
QY 421 HGAYLGTVTVSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSLTHOAMRGQOIGSYF 480  
DB 352 HGAYLGTVTVSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSLTHOAMRGQOIGSYF 411  
QY 481 GSEITSDIDGQVTDVLLVGAPMYFNEGRGRGVYVYELQNRFPVNGTLKDSHESYONA 540  
DB 412 GSEITSDIDGQVTDVLLVGAPMYFNEGRGRGVYVYELQNRFPVNGTLKDSHESYONA 471  
QY 541 RFGSSIASVRDLNODSYNDVNVGAPLEDNHAGAIYIFHFGSGSLTKPKORITASELATG 600  
DB 472 RFGSSIASVRDLNODSYNDVNVGAPLEDNHAGAIYIFHFGSGSLTKPKORITASELATG 531  
QY 601 LOYFGCSIHGQLDNLNEDGLIDLAVGALGNVILASRPVYVQINASHLHFEPSKINIFHRDCK 660  
DB 532 LOYFGCSIHGQLDNLNEDGLIDLAVGALGNVILASRPVYVQINASHLHFEPSKINIFHRDCK 591  
QY 661 RSGRDATCLAAFLCPTPIFLAPHFQTTVIGIRYNATMDERYTPRAHLDGSGDRFTNRVAV 720  
DB 592 RSGRDATCLAAFLCPTPIFLAPHFQTTVIGIRYNATMDERYTPRAHLDGSGDRFTNRVAV 651  
QY 721 LLSQGLCERINPFLVDTADYVKPVTFSVYSLEDPDHGMDDGWPTTLRVSVFPWNG 780  
DB 652 LLSQGLCERINPFLVDTADYVKPVTFSVYSLEDPDHGMDDGWPTTLRVSVFPWNG 711  
QY 781 CNEDEHCVDPDLVDARSDLPTAMEYCORVLRKPAQDCSAYTSLFDTTTFVFIESTQRVAV 840  
DB 712 CNEDEHCVDPDLVDARSDLPTAMEYCORVLRKPAQDCSAYTSLFDTTTFVFIESTQRVAV 771  
QY 841 EATLNRCENAVSTVLNTSOSANLQFASLIQKESDGSIECVNEERLQKQCNVSPFPF 900  
DB 772 EATLNRCENAVSTVLNTSOSANLQFASLIQKESDGSIECVNEERLQKQCNVSPFPF 831  
QY 901 RAKAVAFRLDSFSSKSIPLHLEIELAAGSDNSNERDSTKEDNVAPLRFHLKYEADVLFT 960  
DB -----

Db 832 RAKAVAFRLDSFSSKSIPLHLEIELAAGSDNSNERDSTKEDNVAPLRFHLKYEADVLFT 891  
QY 961 RSSSLSHYEVKLNSSLERYDGIQPPESCIFRIQNLGLPFIHGMMKTIPIATSGNRLL 1020  
Db 892 RSSSLSHYEVKLNSSLERYDGIQPPESCIFRIQNLGLPFIHGMMKTIPIATSGNRLL 951  
QY 1021 KLRFPLTDS-ANTSCNINMGSTVYRPPVVEDLERRAPQLMHSNSDVVSNICNIELVPOE 1079  
Db 952 KLRFPLTDS-ANTSCNINMGSTVYRPPVVEDLERRAPQLMHSNSDVVSNICNIELVPOE 1011  
QY 1080 INPHLLGNLWLSLKALKYKSMKIMVNAALQROFHSPPFIREDPSPSQIIFEISKQBDWQ 1139  
Db 1012 INPHLLGNLWLSLKALKYKSMKIMVNAALQROFHSPPFIREDPSPSQIIFEISKQBDWQ 1071  
QY 1140 VPIWIIIVGSTLGGILLALLVLAIRKLGFPSARRRREPGLDTPPKYLE 1188  
Db 1072 VPIWIIIVGSTLGGILLALLVLAIRKLGFPSARRRREPGLDTPPKYLE 1120  
RESULT 14  
AAB50087  
ID AAB50087 standard; protein; 1188 AA.  
XX  
XX AAB50087;  
XX AC  
XX XX  
DT 19-MAR-2001 (first entry)  
XX  
DE Murine A259.  
XX  
KW Human; A259; integrin alpha subunit; fibrosis; liver disease; lung;  
KW kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;  
KW rheumatoid arthritis.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT Domain 1..1141  
FT Peptide 1..22 /label= Extracellular\_domain  
FT Protein 23..1188 /label= Signal\_peptide  
FT Domain 39..74 /label= Mature\_protein  
FT Domain /label= Integrin\_alphasubunit\_repeat\_domain\_#1  
FT Domain 115..157 /label= Integrin\_alphasubunit\_repeat\_domain\_#2  
FT Domain 164..345 /label= I domain  
FT Domain 367..392 /label= Integrin\_alphasubunit\_repeat\_domain\_#3  
FT Domain 421..455 /label= Integrin\_alphasubunit\_repeat\_domain\_#4  
FT Domain 478..516 /label= Integrin\_alphasubunit\_repeat\_domain\_#5  
FT Domain 540..575 /label= Integrin\_alphasubunit\_repeat\_domain\_#6  
FT Domain 602..640 /label= Integrin\_alphasubunit\_repeat\_domain\_#7  
FT Domain 1142..1164 /label= Transmembrane\_domain  
FT Domain 1165..1188 /label= Cytoplasmic\_domain  
XX WO200073339-A1.  
XX  
XX 07-DEC-2000.  
XX  
XX 15-MAY-2000; 2000WO-US013262.  
XX  
XX 28-MAY-1999; 99US-00322790.  
XX 27-APR-2000; 2000US-00561263.  
XX (MILL-) MILLENNIUM PHARM INC.  
PA

|    |  |  |
|----|--|--|
| XX | Pan Y, Lora JM;  |  |
| XX | WPI; 2001-041142/05.   |  |
| DR | N-PSDB; AAC91904, AAC91905.  |  |
| XX |  |  |
| PT | Nucleic acid encoding alpha-integrin subunits, useful for treatment and diagnosis of fibrosis, e.g. of the liver.  |  |
| PT |  |  |
| XX |  |  |
| PS | Claim 8; Fig 5; 164pp; English.  |  |
| XX |  |  |
| CC | The present sequence is murine integrin alpha subunit, A259 is homologous with the alpha1 and alpha2 integrin subunits and is overexpressed in fibrosis. A259 is implicated in regulation of proliferation, differentiation and/or function of many different cell types. Inhibitors of A259 activity are useful for the treatment of liver disease, particularly fibrosis, and also fibrosis in other organs (specifically lung and kidney). In addition, A259 can be used for treatment and prevention of cancer, osteoporosis, acute myeloid leukaemia, HIV infection, and rheumatoid arthritis |  |
| XX |  |  |
| SQ | Sequence 1188 AA;  |  |
|    | Query Match 91.5%; Score 5693; DB 4; Length 1188;  |  |
|    | Best Local Similarity 90.2%; Pred. No. 0;  |  |
|    | Matches 1071; Conservative 57; Mismatches 60; Indels 0; Gaps 0;  |  |
| QY | 1 MDLPRGLVAVWALSLMPGFTDTFNMDTRKPRVTPGSRRTAFPGYTVQOQHDISGNKMLVWGA 60  |  |
| DB | 1 MDPFRGLVAVWTLSLMPGFTDTFNMDTRNPRVTPGSAAPFGYTVQOQHDISGNKMLVWGA 60  |  |
| QY | 61 PLBTNGVQKTDYKCPVTHGNTKLNGLVTLNVSERKDNRLGLSLATNPKNLSFLA 120  |  |
| DB | 61 PMETNGHQTKDYKCPVTCQNTKLNGLVTLNVSERKDNRLGLSLATNPKNLSFLA 120  |  |
| QY | 121 CSPLWSEHCSSYTTGMSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWYE 180   |  |
| DB | 121 CSPLWSEHCSSYTTGMSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWYE 180   |  |
| QY | 181 VQHLNLIKVFYIGQIQVGVVQYGEDVWHFHLNDRSVKDVVAAASHIEQRGTET 240  |  |
| DB | 181 VQHLNLIKVFYIGQIQVGVVQYGEDVWHFHLNDRSVKDVVAAASHIEQRGTET 240  |  |
| QY | 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGHSDSPLEKVIQOSERDNNVTRAVAVL 300   |  |
| DB | 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGHSDSPLEKVIQOSERDNNVTRAVAVL 300   |  |
| QY | 301 GYNNRGINPETFLNEIKYIASDDKHFFNVTDAAALKDIDVALGDRIFPSLEGTKNET 360  |  |
| DB | 301 GYNNRGINPETFLNEIKYIASDDKHFFNVTDAAALKDIDVALGDRIFPSLEGTKNET 360  |  |
| QY | 361 SPGLEMSQTFSSHVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPPELKN 420  |  |
| DB | 361 SPGLEMSQTFSSHVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPPELKN 420  |  |
| QY | 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNHTGKVIPLTMNNRSLTIHQAMRQOQIGSYF 480  |  |
| DB | 421 HAAVLGYTVTSVSSRQGRVYVAGAPRNHTGKVIPLTMNNRSLTIHQALRGEIQIGSYF 480   |  |
| QY | 481 GSITSDIDGDGVTDVLLGAPWYFSEGRGKVVYELRQKRFVYNGTLDKSHSVQNA 540   |  |
| DB | 481 GSITSDVNDVDRVTDVLLGAPWYFSEGRGKVVYVYNNRQRFVYNGTLDKSHSVQNA 540   |  |
| QY | 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIPHGFRGSILKTPKQRTITASELATG 600   |  |
| DB | 541 RFGSCIASVQDLNQDSYNDVVGAPLEDNHRGAIYIPHGFTNLKPKMQRTITASELAPG 600   |  |
| QY | 601 LQYFGCSIHGOLDLNEGLDLDLAVGALGNVILMSRPVQINASLHPEPSKINIHRDCK 660  |  |
| DB | 601 LQHFPGCSIHGOLDLNEGLDLDLAVGALGNVILMSRPVQINASLHPEPSKINIHRDCK 660   |  |
| QY | 661 RSGRDATCLAAFLCFTPIFLAPHEQTATVGRYNAWDERRYMPRAHLDEGDOFTNRAV 720  |  |

|           |  |  |
|-----------|--|--|
| DB        | 661 RNRGRDATCLAAFLCFTPIFLAPHEQTATVGRYNAWDERRYMPRAHLDEGDOFTNRAV 720   |  |
| QY        | 721 LLSGQELCERINFHVLDTADYVKPVTFVSEYSLDDPHGEMLDGDPWFTTLRVSPVPWNG 780  |  |
| DB        | 721 LLSGQEHQCRINFHVLDTADYVKPVTFVSEYSLDDPHGEMLDGDPWFTTLRVSPVPWNG 780  |  |
| QY        | 781 CNEDEHCVDPDVLVDARSDLPTAMEYCORVLKPKAQDCSAYTSLSDFTTVPFIIESTRQYAV 840   |  |
| DB        | 781 CNEDEHCVDPDVLVDARSDLPTAMEYCORVLKPKAQDCSSYTLSDFTTVPFIIESTRRYAV 840  |  |
| QY        | 841 EATLENRGENAYSTVLANISQANLQFASLIQKEDSDGSIECVNBERRLKQKQVNCVSYPPF 900  |  |
| DB        | 841 EATLENRGENAYSAVLNISQENLQFASLIQKEDSDNSIECVNBERRLKVKVNCVSYPPF 900  |  |
| QY        | 901 RAKAVAFRLDSEFSKSIPLHLIELAAGSDSNERDSTKEDNVAPLPHLKYEADVLFT 960   |  |
| DB        | 901 RAKAVAFRLDSEFSKSVFLHLQLHLAGSDSHEQDSTADDNTALLPHLKYEADVLFT 960   |  |
| QY        | 961 RSSLSHYEVKLSLEYDYGIGPPSPCIPFQNLGLPPIHGMWKIPIPIATSGNELL 1020  |  |
| DB        | 961 RSSLSHFVANSLSLEYDYGIGPPSPCVFQNLGLPPIHGMWKIPIPIATSGNELL 1020  |  |
| QY        | 1021 KLRDPLTDEANTSCNMGNSTEYRTPVEEDLRAPQLNHSNDSVVSNCNIRLVNPEI 1080  |  |
| DB        | 1021 MURDFTDQNTSCNMGNSTEYRSTPTEEDLSHAPORNHSNDSVVSICNLRAPSQET 1080  |  |
| QY        | 1081 NPHLLGNLWLSLAKYKSMKIMVNAALQRFHSPIFREEDPSRQIEFISKOEDMQV 1140   |  |
| DB        | 1081 SPYLVGNLWLSLAKYKSLKITVNAALQRFHSPIFREEDPSRQIEFISKOEDMQV 1140   |  |
| QY        | 1141 PIWIIIVGSTLGLLALLLALLLALRKLGFFRSARRRREFGLDPTPKVLE 1188  |  |
| DB        | 1141 PIWIIIVGSTLGLLALLLALLLALWKLGFFRSARRRREFGLDPTPKVLE 1188  |  |
| RESULT 15 |  |  |
| AAU10552  |  |  |
| ID        | AAU10552 standard; protein; 1188 AA.   |  |
| XX        | AAU10552;  |  |
| AC        |  |  |
| XX        |  |  |
| DT        | 14-FEB-2002 (first entry)  |  |
| XX        |  |  |
| DE        | Murine A259 polypeptide.   |  |
| KW        | Mouse; A259; integrin alpha subunit; integrin alpha 10; secreted protein; liver disease; fibrosis; lung; kidney; bone associated disorder; blood; cartilage associated disorder; haematopoietic disorder; bone marrow; immune related disease; apoptotic disorder; neuronal tissue disease; neurodegenerative disease; gene therapy; cancer; cytostatic; osteopathic; nephrotropic; immunomodulator; anti-inflammatory; neuroprotective; antiarthritic; antianemic; antiallergic; antiasthmatic; dermatological; antidiabetic; anticonvulsant; antiparkinsonian. |  |
| XX        |  |  |
| OS        | Mus musculus.  |  |
| XX        |  |  |
| Key       | Location/Qualifiers  |  |
| FT        | Domain 1..1141 /note= "Extracellular domain"   |  |
| FT        | Peptide 1..22 /note= "Signal peptide"  |  |
| FT        | Protein 23..1188 /note= "Mature murine A259"   |  |
| FT        | Domain 39..74 /note= "Integrin alpha repeat domain"  |  |
| FT        | Domain 115..157 /note= "Integrin alpha repeat domain"  |  |
| FT        | Domain 164..345 /note= "I domain or Von Willebrand Factor type A domain"   |  |
| FT        | Domain 367..392 /note= "Integrin alpha repeat domain"  |  |
| FT        | Domain 421..455 /note= "Integrin alpha repeat domain"  |  |
| FT        |  |  |

FT Domain 478. .516  
 FT /note= "Integrin alpha repeat domain"  
 FT Domain 540. .575  
 FT /note= "Integrin alpha repeat domain"  
 FT Domain 602. .640  
 FT /note= "Integrin alpha repeat domain"  
 FT Domain 1142. .1164  
 FT /note= "Transmembrane domain"  
 FT Domain 1165. .1188  
 FT /note= "Cytoplasmic domain"  
 XX  
 PN W0200181414-R2.  
 XX  
 XX 01-NOV-2001.  
 PD  
 XX  
 XX 27-APR-2001; 2001WO-US013516.  
 PR  
 XX 27-APR-2000; 2000US-00561263.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 PI  
 PI Pan Y, Lora J;  
 XX  
 DR WPI; 2002-041397/05.  
 DR N-PSDB; AAS16874.  
 XX  
 PT New A259 nucleic acids and polypeptides, which comprise integrin alpha  
 PT subunit, useful for diagnosing, preventing or treating e.g. liver  
 PT disease, kidney or lung fibrosis, cancers, blood disorders or immune  
 PT related diseases.  
 XX  
 PS Claim 9; Fig 5; 168pp; English.  
 XX  
 CC The invention relates to human and murine A259 nucleic acid molecules  
 CC which encode secreted proteins with homology to integrin alpha subunits,  
 CC specifically to integrin alpha 10. The A259 polypeptide and nucleic acid  
 CC are useful for treating liver disease or fibrosis, particularly kidney  
 CC fibrosis or lung fibrosis. The A259 polypeptide and nucleic acid are also  
 CC useful for diagnosing, preventing or treating cartilage and bone  
 CC associated disorders (such as bone cancer, achondroplasia, myeloma,  
 CC fibrous dysplasia, scoliosis, osteoarthritis, osteosarcoma and  
 CC osteoporosis), bone marrow, blood and haematopoietic disorders (such as  
 CC acute myeloid leukaemia, haemophilia, anaemia and thalassemia), immune  
 CC related diseases (such as AIDS, viral infections, cancers, T cell  
 CC autoimmune disorders e.g. HIV, and allergic inflammatory disorders e.g.  
 CC asthma and psoriasis), apoptotic disorders (such as systemic lupus  
 CC erythematosus and insulin-dependent diabetes mellitus), diseases of the  
 CC neuronal tissues (such as epilepsy and muscular dystrophy) and  
 CC neurodegenerative diseases (such as Parkinson's disease and Huntington's  
 CC disease). This sequence represents the murine A259 polypeptide  
 XX  
 SQ Sequence 1188 AA;  
 Query Match 91.5%; Score 5693; DB 5; Length 1188;  
 Best Local Similarity 90.2%; Fred. No. 0;  
 Matches 1071; Conservative 57; Mismatches 60; Indels 0; Gaps 0;  
 QY 1 MDLPRLVVAWALSIPGPTDFNMNDRKPRVPGSRFAFGYTVQOHDISGNKWL VVGA 60  
 DB 1 MDPPRLVAVNTLSLPWPTDFNMNDRKPRVPGSRFAFGYTVQOHDISGNKWL VVGA 60  
 QY 61 PLETNGYKGTGVYKCPVHNGCTKLNLRUTLSNVSRKKNMRLGLSLATNPKNSPLA 120  
 DB 61 PMETNGHQKGTGVYKCPVQGNCTKLNLRUTLSNVSRKKNMRLGLSLATNPKNSPLA 120  
 QY 121 CSPLWSHCGSGSYTTGCMRSVNSNFRSKTVAPALQRCQTYMDIVILVDCNSNTPWVE 180  
 DB 121 CSPLWSHCGSGSYTTGCMRSVNSNFRSKTVAPALQRCQTYMDIVILVDCNSNTPWVE 180  
 QY 181 VOHFLINILKKFYIGPGIYGVQYGDVHVEHFLNDRSVKGVVVEAASHIEQRGGTET 240  
 DB 181 VOHFLINILKKFYIGPGIYGVQYGDVHVEHFLNDRSVKGVVVEAASHIEQRGGTET 240

Search completed: June 24, 2004, 17:54:32  
 Job time : 71 secs

QY 241 RTAFGIEFARSEAFQKGRGAKVMIVITDGSBHSBSPDLEKVIQOSERONVTRYAVAVL 300  
 DB RTAFGIEFARSEAFQKGRGAKVMIVITDGSBHSBSPDLEKVIQOSERONVTRYAVAVL 300  
 QY 301 GYNRRGINPETFELNEIKYIASDPDDKHFFNVVDEAALKDIVDALGDRIFSLGKTNKNET 360  
 DB GYNRRGINPETFELNEIKYIASDPDDKHFFNVVDEAALKDIVDALGDRIFSLGKTNKNET 360  
 QY 361 SFGLEMSQTGFSSHVVEDGILLGAVGAYDWNAGVAVLKETSAGKVIPLRESYLYKEPPEBLKN 420  
 DB SFGLEMSQTGFSSHVVEDGILLGAVGAYDWNAGVAVLKETSAGKVIPLRESYLYKEPPEBLKN 420  
 QY 421 HGAYLYGTVTSVSVSSRQGRVYVAGAPRNFHTKGVILFTMNNESLTHQAMRGQIGSYF 480  
 DB HAAYLYGTVTSVSVSSRQGRVYVAGAPRNFHTKGVILFTMNNESLTHQAMRGQIGSYF 480  
 QY 481 GSEITSDIDGDGVDVLLVAGPMYFNEGRERKGVYVYELRQNRFFVYNGTLKSHSYQNA 540  
 DB GSEITSDVVDNDVDRVTDVLLVAGPMYFSEGRERKGVYVYELRQNRFFVYNGTLKSHSYQNA 540  
 QY 541 RFGSSIASVDLNDQSDYNVDVVGAPLEDNHAGAIYIFHGFSGSTLTKPKORITASELATG 600  
 DB RFGSCIASVDLNDQSDYNVDVVGAPLEDNHRGAIYIFHGFQTNILKKPKORITASELATG 600  
 QY 601 LOYFGCSIHGOLDINEDGLIDLAVGALGNVAVILWSRPVQVQINASLHFEPSKINIFHRDCK 660  
 DB LOHFGCSIHGOLDINEDGLVDLAVGALGNVAVILWSRPVQVQINASLHFEPSKINIFHRDCK 660  
 QY 661 RSGRDATCLAAFLCFTPIFLAPHFTTUGIRINATMDERRYTPRAHLDEGGDFTRAV 720  
 DB RSGRDATCLAAFLCFTPIFLAPHFTTUGIRINATMDERRYTPRAHLDEGGDFTRAV 720  
 QY 721 LLSSGGQELCRINFHVLDTADYVYKPTVFSVEYSLEDPDHGMPLDDGPTTLRVSVFPWNG 780  
 DB LLSSGGQELCRINFHVLDTADYVYKPAVSVEYSLEDPDHGMPLDDGPTTLRVSVFPWNG 780  
 QY 781 CNEDEHCVPLDLVARSDDLPTAMEYCORVLKPAQDCSAYTLSPDTTVFIIESTRQVAV 840  
 DB CNEDEHCVPLDLVARSDDLPTAMEYCORVLKPAQDCSAYTLSPDTTVFIIESTRQVAV 840  
 QY 841 EATLENRGENAYSTVLNISOSANIQFASLICKEDSDGSIKCVNEERLQKOVNVSYPFF 900  
 DB EATLENRGENAYSAVLNISQSENLQFASLICKEDSDGSIKCVNEERLQKOVNVSYPFF 900  
 QY 901 RAKAKVAFRLDSEFSSKIFLHLEIEILAAGSDSNRDKEDNVAPLRFHLKYEADVLT 960  
 DB RAKAKVAFRLDSEFSSKIFLHLEIEILAAGSDSHEQSDTADDNTALLRFLHLYEADVLT 960  
 QY 961 RSSSLSHYVKLNSLERYDGIQPPPSCTRIQNHGLFPIHGMKMTIPIATESGNRL 1020  
 DB RSSSLSHYVKLNSLERYDGIQPPPSCTRIQNHGLFPIHGMKMTIPIATESGNRL 1020  
 QY 1021 KLDRDLTDEANTSCNINWNSNTEYRPTVEEDLRRAPQLNHSNSDVWSINCNIRLVPQEI 1080  
 DB MLRDFDFTDQNTSCNINWNSNTEYRPTVEEDLRRAPQLNHSNSDVWSINCNIRLVPQEI 1080  
 QY 1081 NFHLNGLMLRSALKYKSKMIMVNAALQORPHSPPIFREEDPSRQIEFISKQEDQV 1140  
 DB SFYLVGNLMLTSLKALKYRSKITVNAALQORPHSPPIFREEDPSRQIEFISKQEDQV 1140  
 QY 1141 PIMTIIVGSTIGGLLALLVALLKLGKFFRSARRRRPGLOPTPKVLE 1188  
 DB PIMTIIVGSTIGGLLALLVALLKLGKFFRSARRRRPGLOPTPKVLE 1188

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 17:56:16 ; Search time 54 Seconds  
(without alignments)  
6210.900 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 6224

Sequence: 1 MDLPRGLVWALSLWPGFT.....FRSARRRRPGLDTPKYLE 1188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1153542 seqs, 282313645 residues

Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09D\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10F\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match % | Length | DB ID | Description         |
|------------|--------|---------------|--------|-------|---------------------|
| 1          | 6196   | 99.6          | 1188   | 15    | US-10-291-265-338   |
| 2          | 6192   | 99.5          | 1188   | 15    | US-10-291-265-810   |
| 3          | 6188.5 | 99.4          | 1189   | 10    | US-09-984-130-35    |
| 4          | 6188.5 | 99.4          | 1189   | 10    | US-09-984-130-35    |
| 5          | 6188.5 | 99.4          | 1189   | 12    | US-10-262-839-4     |
| 6          | 5780   | 92.9          | 1120   | 12    | US-10-262-839-6     |
| 7          | 5393   | 86.5          | 1034   | 10    | US-09-984-130-43    |
| 8          | 5393   | 86.5          | 1034   | 10    | US-09-984-130-43    |
| 9          | 3621   | 58.2          | 707    | 9     | US-09-764-870-313   |
| 10         | 3621   | 58.2          | 707    | 14    | US-10-125-540-313   |
| 11         | 3182   | 51.1          | 688    | 10    | US-09-866-050A-624  |
| 12         | 3113   | 50.0          | 696    | 10    | US-09-866-050A-501  |
| 13         | 2518   | 40.5          | 545    | 10    | US-09-866-050A-500  |
| 14         | 2200   | 35.3          | 437    | 15    | US-10-108-260A-3386 |
| 15         | 2166.5 | 34.8          | 1179   | 12    | US-09-918-715-250   |

|    |        |      |      |    |                    |                    |
|----|--------|------|------|----|--------------------|--------------------|
| 16 | 2159   | 34.7 | 1151 | 10 | US-09-984-130-103  | Sequence 103, App  |
| 17 | 2159   | 34.7 | 1151 | 10 | US-09-984-130-103  | Sequence 103, App  |
| 18 | 2149.5 | 34.5 | 1180 | 12 | US-09-918-715-307  | Sequence 107, App  |
| 19 | 1863   | 29.9 | 1181 | 12 | US-10-211-463-187  | Sequence 187, App  |
| 20 | 1863   | 29.9 | 1181 | 14 | US-10-160-354-2    | Sequence 2, Appli  |
| 21 | 1863   | 29.9 | 1181 | 15 | US-10-295-027-1286 | Sequence 1286, Ap  |
| 22 | 1809   | 29.1 | 1147 | 12 | US-10-336-603A-42  | Sequence 42, Appli |
| 23 | 1194.5 | 19.2 | 1161 | 9  | US-09-350-259-53   | Sequence 53, Appli |
| 24 | 1194.5 | 19.2 | 1161 | 10 | US-09-891-943-53   | Sequence 53, Appli |
| 25 | 1184.5 | 19.0 | 1161 | 9  | US-09-350-259-55   | Sequence 55, Appli |
| 26 | 1184.5 | 19.0 | 1161 | 10 | US-09-891-943-55   | Sequence 55, Appli |
| 27 | 1181.5 | 19.0 | 1155 | 9  | US-09-350-259-46   | Sequence 46, Appli |
| 28 | 1181.5 | 19.0 | 1155 | 10 | US-09-891-943-46   | Sequence 46, Appli |
| 29 | 1167.5 | 18.8 | 1151 | 9  | US-09-350-259-37   | Sequence 37, Appli |
| 30 | 1167.5 | 18.8 | 1151 | 10 | US-09-891-943-37   | Sequence 37, Appli |
| 31 | 1146.5 | 18.4 | 1170 | 9  | US-09-945-265-2    | Sequence 2, Appli  |
| 32 | 1146.5 | 18.4 | 1170 | 12 | US-10-261-164-1    | Sequence 1, Appli  |
| 33 | 1142.5 | 18.4 | 1161 | 9  | US-09-350-259-99   | Sequence 99, Appli |
| 34 | 1142.5 | 18.4 | 1161 | 10 | US-09-891-943-99   | Sequence 99, Appli |
| 35 | 1139   | 18.3 | 1161 | 9  | US-09-350-259-2    | Sequence 2, Appli  |
| 36 | 1139   | 18.3 | 1161 | 10 | US-09-891-943-2    | Sequence 2, Appli  |
| 37 | 1136   | 18.3 | 1223 | 16 | US-10-408-765A-295 | Sequence 295, App  |
| 38 | 1132.5 | 18.2 | 1163 | 15 | US-10-116-275-204  | Sequence 204, App  |
| 39 | 1103.5 | 17.7 | 1163 | 9  | US-09-350-259-4    | Sequence 4, Appli  |
| 40 | 1103.5 | 17.7 | 1163 | 10 | US-09-891-943-4    | Sequence 4, Appli  |
| 41 | 1094   | 17.6 | 1153 | 9  | US-09-350-259-3    | Sequence 3, Appli  |
| 42 | 1094   | 17.6 | 1153 | 10 | US-09-302-481A-1   | Sequence 1, Appli  |
| 43 | 1094   | 17.6 | 1153 | 10 | US-09-891-943-3    | Sequence 3, Appli  |
| 44 | 1094   | 17.6 | 1153 | 14 | US-10-144-259-30   | Sequence 30, Appli |
| 45 | 1094   | 17.6 | 1153 | 14 | US-10-207-655-176  | Sequence 176, App  |

## ALIGNMENTS

### RESULT 1

US-10-291-265-338  
; Sequence 338, Application US/10291365  
; Publication No. US20030232054A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; APPLICANT: Tang et al  
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-017 (785)  
; CURRENT APPLICATION NUMBER: US/10/291,265  
; CURRENT FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/491,404  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/617,746  
; PRIOR FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 09/631,451  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 09/633,870  
; PRIOR FILING DATE: 2000-09-15  
; NUMBER OF SEQ ID NOS: 944  
; SOFTWARE: fastseq for Windows Version 3.0  
; SEQ ID NO 338  
; LENGTH: 1188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-291-265-338

Query Match 99.6%; Score 6196; DB 15; Length 1188;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1184; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

|    |    |  |     |
|----|----|--|-----|
| Qy | 1  | MDLPRGLVWALSLWPGFTDTFNNDYRKPRVIGSRTAFPGYTVQGHDSGNKWLWVGA | 60  |
| Db | 1  | MDLPRGLVWALSLWPGFTDTFNNDYRKPRVIGSRTAFPGYTVQGHDSGNKWLWVGA | 60  |
| Qy | 61 | PLETNGYKTDGVYKCPVHGNCTKLNLRVTLSNVSRKDNRLGLSLATNPKNDSFLA  | 120 |
| Db | 61 | PLETNGYKTDGVYKCPVHGNCTKLNLRVTLSNVSRKDNRLGLSLATNPKNDSFLA  | 120 |

QY 121 CSPWSECGSSYTTGCMCSRVSNSRPSKTVAPALORCQTYMDIVIVLDGNSIYPWVE 180  
DB |||||  
QY 121 CSPWSECGSSYTTGCMCSRVSNSRPSKTVAPALORCQTYMDIVIVLDGNSIYPWVE 180  
DB |||||  
QY 181 VOHFLINILKKFYIGPGQIQGVVQYGEDVYHFFHLDYASVDVVEAAASHIEQGGTET 240  
DB |||||  
QY 241 RTAPGIEFARSEAFQKGRKGAKKVMIVITDGSNDSPLDKVIOQSERDNTRYAVAVL 300  
DB |||||  
QY 301 GYNNRGINPETFLNEIKYIASDDPKHFFNVTDAAALDXIDVLDGDRIFSLGTTNKNET 360  
DB |||||  
QY 361 SFGLMSQTFSSHVVEDGVLGAVGYDNGAVLKETSAGKVIPLRESYLKEPPELKN 420  
DB |||||  
QY 421 HGAYLGYTTSVSSRQGRVYVAGAPRNFHTGKVIPLFTMNNRSLTIHQAMRQOIGSYF 480  
DB |||||  
QY 481 GSEITSDVDIDGQVTDVLLVGAAPMYFNEGERGKVVYVELRQNRFYNGTLKDSHSYQNA 540  
DB |||||  
QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFPGSILKTPQRITASELANG 600  
DB |||||  
QY 601 LQYFGCSIHGQDLNEDGLDLAGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660  
DB |||||  
QY 661 RSGRDATCLAAFCPTPIFLAPHFQTTVCGIRNATMDERRYTTPRAHLDGGRDPTNRVAV 720  
DB |||||  
QY 721 LLSGGQELCERINPHVLDIADYVKPTFSVEYSLEDPDHGEMDGMPTTLRVSPFPWNG 780  
DB |||||  
QY 781 CNEDEHCVPDLVLDARSIDUPTAMEYCORVLRKPAQDCSAYTISFDTTVEIIESTRQRAV 840  
DB |||||  
QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERLQKQCNVSPFP 900  
DB |||||  
QY 901 RAKAKVAFRLDSEFSKIFLHLEIELAAGSDSNERDSTKEDNVAPLRFLKYEADVLPT 960  
DB |||||  
QY 961 RSSLSHYEVKLNSSLERYDGIQPPSCIFRIONLGLPFIHGMKMITIPIATRSNRL 1020  
DB |||||  
QY 1021 KLDRDFTDRANTSCNIWGSTETPTPVEEDLERAPQLNHSNDSVVSINCNRLVFNQBI 1080  
DB |||||  
QY 1081 NFHLNGLNLSKALKYSKIMVNAALQROPHSPFIIFREEDPSRQIEFEISKQEDMOV 1140  
DB |||||  
QY 1141 PIWIIWGSLGGLLALVLAALRKLGFPPRSARRRREPLDPTPKYLE 1188  
DB |||||

RESULT 2  
US-10-291-265-810  
; Sequence 810, Application US/10291265  
; Publication No. US20030232054A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; APPLICANT: Tang et al  
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-017 (785)  
; CURRENT APPLICATION NUMBER: US/10/291,265  
; CURRENT FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/491,404  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/617,746  
; PRIOR FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 09/631,451  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 09/633,870  
; PRIOR FILING DATE: 2000-09-15  
; NUMBER OF SEQ ID NOS: 944  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 810  
; LENGTH: 1188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-291-265-810  
  
Query Match 99.5%; Score 6192; DB 15; Length 1188;  
Best Local Similarity 99.6%; Pred No. 0;  
Matches 1183; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
QY 1 MDLPRGLVAVALSILWPGFTDTFNNMTRKPRVIGSRVTAFFGYTVOQHDISGNKMLVYGA 60  
DB |||||  
QY 1 MDLPRGLVAVALSILWPGFTDTFNNMTRKPRVIGSRVTAFFGYTVOQHDISGNKMLVYGA 60  
DB |||||  
QY 61 PLENTGYOKTGDVVKCPVIHGNCTKLNLRVYTLNVSEKKNMRGLSLATNPKNDSFLA 120  
DB |||||  
QY 61 PLENTGYOKTGDVVKCPVIHGNCTKLNLRVYTLNVSEKKNMRGLSLATNPKNDSFLA 120  
DB |||||  
QY 121 CSPWSECGSSYTTGCMCSRVSNSRPSKTVAPALORCQTYMDIVIVLDGNSIYPWVE 180  
DB |||||  
QY 121 CSPWSECGSSYTTGCMCSRVSNSRPSKTVAPALORCQTYMDIVIVLDGNSIYPWVE 180  
DB |||||  
QY 181 VOHFLINILKKFYIGPGQIQGVVQYGEDVYHFFHLDYASVDVVEAAASHIEQGGTET 240  
DB |||||  
QY 241 RTAPGIEFARSEAFQKGRKGAKKVMIVITDGSNDSPLDKVIOQSERDNTRYAVAVL 300  
DB |||||  
QY 241 RTAPGIEFARSEAFQKGRKGAKKVMIVITDGSNDSPLDKVIOQSERDNTRYAVAVL 300  
DB |||||  
QY 301 GYNNRGINPETFLNEIKYIASDDPKHFFNVTDAAALDXIDVLDGDRIFSLGTTNKNET 360  
DB |||||  
QY 301 GYNNRGINPETFLNEIKYIASDDPKHFFNVTDAAALDXIDVLDGDRIFSLGTTNKNET 360  
DB |||||  
QY 361 SFGLMSQTFSSHVVEDGVLGAVGYDNGAVLKETSAGKVIPLRESYLKEPPELKN 420  
DB |||||  
QY 361 SFGLMSQTFSSHVVEDGVLGAVGYDNGAVLKETSAGKVIPLRESYLKEPPELKN 420  
DB |||||  
QY 421 HGAYLGYTTSVSSRQGRVYVAGAPRNFHTGKVIPLFTMNNRSLTIHQAMRQOIGSYF 480  
DB |||||  
QY 421 HGAYLGYTTSVSSRQGRVYVAGAPRNFHTGKVIPLFTMNNRSLTIHQAMRQOIGSYF 480  
DB |||||  
QY 481 GSEITSDVDIDGQVTDVLLVGAAPMYFNEGERGKVVYVELRQNRFYNGTLKDSHSYQNA 540  
DB |||||  
QY 481 GSEITSDVDIDGQVTDVLLVGAAPMYFNEGERGKVVYVELRQNRFYNGTLKDSHSYQNA 540  
DB |||||  
QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFPGSILKTPQRITASELANG 600  
DB |||||  
QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFPGSILKTPQRITASELANG 600  
DB |||||  
QY 601 LQYFGCSIHGQDLNEDGLDLAGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660  
DB |||||



Db 601 LQYFGCSIHGQDLNEDGLIDLVAGLGNVILMSRPVQINASHFEPSPKINIFHRDCK 660  
Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVIGIRYNATMDERRYPTRAHLDEGDRFTNRAV 720  
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVIGIRYNATMDERRYPTRAHLDEGDRFTNRAV 720  
Qy 721 LLSGQBELCERINFHVLDADYVVKVTSVEYSLEDDPHGPKLDDGWTTILRVSVPPFWNG 780  
Db 721 LLSGQBELCERINFHVLDADYVVKVTSVEYSLEDDPHGPKLDDGWTTILRVSVPPFWNG 780  
Qy 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLSPDTTTFVFIESTRQVAV 840  
Db 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLSPDTTTFVFIESTRQVAV 840  
Qy 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIHCVNEBRRLQKQVNVSPFF 900  
Db 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIHCVNEBRRLQKQVNVSPFF 900  
Qy 901 RAKAKVAFLDSEFSKSIPLHLEIEILAAGSDNSERDSTKEDNVAPLRFHLKYEADVLT 960  
Db 901 RAKAKVAFLDSEFSKSIPLHLEIEILAAGSDNSERDSTKEDNVAPLRFHLKYEADVLT 960  
Qy 961 RSSLSHYEVKLNSSLERYDGIQPPFSCIFRIQNLGLPFIHGMNMKTIPIATRSNRL 1020  
Db 961 RSSLSHYEVKLNSSLERYDGIQPPFSCIFRIQNLGLPFIHGMNMKTIPIATRSNRL 1020  
Qy 1021 KLDRFLTDEANTSCNIGWNSTEYRPTVEEDLRAPQLNHSNDVVSINCNIRLVNQE 1080  
Db 1021 KLDRFLTDEANTSCNIGWNSTEYRPTVEEDLRAPQLNHSNDVVSINCNIRLVNQE 1080  
Qy 1081 NFHLGNLWLSLKALKYKMKIMVNAALQRFHSPFIREFEDPSRQIEFISKOEDQV 1140  
Db 1081 NFHLGNLWLSLKALKYKMKIMVNAALQRFHSPFIREFEDPSRQIEFISKOEDQV 1140  
Qy 1141 PIWIIVCSITLGLLALLLVILALVKLGFRSARRREREGLOPTPKVLE 1188  
Db 1141 PIWIIVCSITLGLLALLLVILALVKLGFRSARRREREGLOPTPKVLE 1188

RESULT 3  
US-09-984-130-35  
; Sequence 35, Application US/09984130  
; Publication No. US2003005231A1  
; GENERAL INFORMATION:  
; APPLICANT: NI et al.  
; TITLE OF INVENTION: 12 Human Secreted Proteins  
; FILE REFERENCE: PF489P2  
; CURRENT APPLICATION NUMBER: US/09/984,130  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/243,792  
; PRIOR FILING DATE: 2000-10-30  
; PRIOR APPLICATION NUMBER: 09/836,353  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/198,407  
; PRIOR FILING DATE: 2000-04-19  
; PRIOR APPLICATION NUMBER: PCT/US99/25031  
; PRIOR FILING DATE: 1999-10-27  
; PRIOR APPLICATION NUMBER: 60/105,971  
; PRIOR FILING DATE: 1998-10-28  
; NUMBER OF SEQ ID NOS: 149  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 35  
; LENGTH: 1189  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-984-130-35

Query Match 99.4%; Score 6188.5; DB 10; Length 1189;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1184; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
Qy 1 MDLPRGLVAVALSILWPGFTDTFNMDFTRKPRVIPGSRATFFGYTVQOHDISGNKWL VVGA 60

Db 1 MDLPRGLVAVALSILWPGFTDTFNMDFTRKPRVIPGSRATFFGYTVQOHDISGNKWL VVGA 60  
Qy 61 PLETNGYQKTGDVYKCPVHIGNCTKLNGLRVTLNSVGERKDNMFLGLSLATNPKNSFLA 120  
Db 61 PLETNGYQKTGDVYKCPVHIGNCTKLNGLRVTLNSVGERKDNMFLGLSLATNPKNSFLA 120  
Qy 121 CSPLWSEHCSSYYTTGKCSRVSNSNFRPSKTVAPALQRCOTYMDIVILVLDGSNSIYPWVE 180  
Db 121 CSPLWSEHCSSYYTTGKCSRVSNSNFRPSKTVAPALQRCOTYMDIVILVLDGSNSIYPWVE 180  
Qy 181 VQFLINILAKFYIGPQOIQGVVVQYGEDVVFHFLNDYRSVKDVVAAASHIEQGGTET 240  
Db 181 VQFLINILAKFYIGPQOIQGVVVQYGEDVVFHFLNDYRSVKDVVAAASHIEQGGTET 240  
Qy 241 RTAPGIEFARSEAFQKGRGAKKVMIVITDGSNDSPLLEKVOQSRDNVTRYAVAVL 300  
Db 241 RTAPGIEFARSEAFQKGRGAKKVMIVITDGSNDSPLLEKVOQSRDNVTRYAVAVL 300  
Qy 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVITDEAALKDIDVALGDRIIFSLEGTNKNET 360  
Db 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVITDEAALKDIDVALGDRIIFSLEGTNKNET 360  
Qy 361 SFGLEMSQTFSSHVVDGVLGAVGAYDNGAVLKETSAGKVIPLRESYIKEPPEELKN 420  
Db 361 SFGLEMSQTFSSHVVDGVLGAVGAYDNGAVLKETSAGKVIPLRESYIKEPPEELKN 420  
Qy 421 HGAVLGYTTSVSSRQRYVVGAPRENHTKVILFTMNNRSLTIHQMRGQOIGSYF 480  
Db 421 HGAVLGYTTSVSSRQRYVVGAPRENHTKVILFTMNNRSLTIHQMRGQOIGSYF 480  
Qy 481 GSEITSVDIDGQDVTDLVVGAPMYFNEGRGKVVYVELRQNFVYNGTLKDSHSYQNA 540  
Db 481 GSEITSVDIDGQDVTDLVVGAPMYFNEGRGKVVYVELRQNFVYNGTLKDSHSYQNA 540  
Qy 541 RFSSSIASVRLNQDSYNDVVGAPLEDNAGAIYIFHGFRGSLTKPKORITASELATG 600  
Db 541 RFSSSIASVRLNQDSYNDVVGAPLEDNAGAIYIFHGFRGSLTKPKORITASELATG 600  
Qy 601 LQYFGCSIHGQDLNEDGLIDLVAGLGNVILMSRPVQINASHFEPSPKINIFHRDCK 660  
Db 601 LQYFGCSIHGQDLNEDGLIDLVAGLGNVILMSRPVQINASHFEPSPKINIFHRDCK 660  
Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVIGIRYNATMDERRYPTRAHLDEGDRFTNRAV 720  
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVIGIRYNATMDERRYPTRAHLDEGDRFTNRAV 720  
Qy 721 LLSGQBELCERINFHVLDADYVVKVTSVEYSLEDDPHGPKLDDGWTTILRVSVPPFWNG 780  
Db 721 LLSGQBELCERINFHVLDADYVVKVTSVEYSLEDDPHGPKLDDGWTTILRVSVPPFWNG 780  
Qy 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLSPDTTTFVFIESTRQVAV 840  
Db 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLSPDTTTFVFIESTRQVAV 840  
Qy 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIHCVNEERLQKQVNVSPFF 900  
Db 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIHCVNEERLQKQVNVSPFF 900  
Qy 901 RAKAKVAFLDSEFSKSIPLHLEIEILAAGSDNSERDSTKEDNVAPLRFHLKYEADVLT 960  
Db 901 RAKAKVAFLDSEFSKSIPLHLEIEILAAGSDNSERDSTKEDNVAPLRFHLKYEADVLT 960  
Qy 961 RSSLSHYEVKLNSSLERYDGIQPPFSCIFRIQNLGLPFIHGMNMKTIPIATRSNRL 1020  
Db 961 RSSLSHYEVKLNSSLERYDGIQPPFSCIFRIQNLGLPFIHGMNMKTIPIATRSNRL 1020  
Qy 1021 KLDRFLTDEANTSCNIGWNSTEYRPTVEEDLRAPQLNHSNDVVSINCNIRLVNQE 1079  
Db 1021 KLDRFLTDEANTSCNIGWNSTEYRPTVEEDLRAPQLNHSNDVVSINCNIRLVNQE 1080  
Qy 1080 INFHLGNLWLSLKALKYKMKIMVNAALQRFHSPFIREFEDPSRQIEFISKOEDQV 1139

Db 1081 INFHLGNLWLSLKALKYKSMKIMVNAALQORQFHSPFFIFREBDDPSRQIVFISKQEDWQ 1140

Qy 1140 VPIWIIIVGSTLGGILLALLVLALRLKLGPPRSARRRREBGLDTPKYLE 1188

Db 1141 VPIWIIIVGSTLGGILLALLVLALWKLGGFRSARRRREBGLDTPKYLE 1189

RESULT 4

US-09-836-353A-35

/ Sequence 35, Application US/09836353A

/ Publication No. US20030129685A1

/ GENERAL INFORMATION:

/ APPLICANT: Ni et al.

/ TITLE OF INVENTION: 12 Human Secreted Proteins

/ CURRENT APPLICATION NUMBER: US/09/836,353A

/ CURRENT FILING DATE: 2001-04-18

/ PRIOR FILING DATE: 2000-04-19

/ PRIOR APPLICATION NUMBER: PCT/US99/25031

/ PRIOR FILING DATE: 1999-10-27

/ PRIOR APPLICATION NUMBER: 60/105,971

/ PRIOR FILING DATE: 1998-10-28

/ NUMBER OF SEQ ID NOS: 147

/ SOFTWARE: Patent In Ver. 2.0

/ SEQ ID NO 35

/ LENGTH: 1189

/ TYPE: PRT

/ ORGANISM: Homo sapiens

US-09-836-353A-35

Query Match 99.4%; Score 6188.5; DB 10; Length 1189;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1184; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 1 MDLPRLGVAVWALSPLPGFTDTFNMDTRKPRVTPGSRTPAFFGYTVQOCHDISGNKMLVGA 60

Db 1 MDLPRLGVAVWALSPLPGFTDTFNMDTRKPRVTPGSRTPAFFGYTVQOCHDISGNKMLVGA 60

Qy 61 PLETNQYQKTDYKCPVHIGNCTKLNLRVLSNVSRKDNRLGSLATNPKNFLA 120

Db 61 PLETNQYQKTDYKCPVHIGNCTKLNLRVLSNVSRKDNRLGSLATNPKNFLA 120

Qy 121 CSPLWSEHCSSYTTTGKMSRVNSNFRSKTVAPALQRCOTYMDIVIVLDGNSIYPWYE 180

Db 121 CSPLWSEHCSSYTTTGKMSRVNSNFRSKTVAPALQRCOTYMDIVIVLDGNSIYPWYE 180

Qy 181 VOHFLINILKFFYIGPQIQGVQYGEDVVEHFLHNDYRSKDVVEAAASHIEQGGTET 240

Db 181 VOHFLINILKFFYIGPQIQGVQYGEDVVEHFLHNDYRSKDVVEAAASHIEQGGTET 240

Qy 241 RTAPGIEFARSEAPQGRGAKKUMIVITDGHSDSPLEKVIQOSERDNVTRAVAVL 300

Db 241 RTAPGIEFARSEAPQGRGAKKUMIVITDGHSDSPLEKVIQOSERDNVTRAVAVL 300

Qy 301 GYNNRRGINPETFLNEIKYIASDDPKHPNVNVDAAALXDIVDALGDRIFSLEGTKNET 360

Db 301 GYNNRRGINPETFLNEIKYIASDDPKHPNVNVDAAALXDIVDALGDRIFSLEGTKNET 360

Qy 361 SPGLEMSQTFSEHVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPELKN 420

Db 361 SPGLEMSQTFSEHVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPELKN 420

Qy 421 HGAYLGVTYSVSSRQGRVYVAGAPRPNHTGKVLFTWNNRSLTIHOMRGOQIGSYF 480

Db 421 HGAYLGVTYSVSSRQGRVYVAGAPRPNHTGKVLFTWNNRSLTIHOMRGOQIGSYF 480

Qy 481 GSBITSVIDDGDVTDVLLVGAQPMYFNEGRGKVVYVELQNRVYNGTLKDSHSYQNA 540

Db 481 GSBITSVIDDGDVTDVLLVGAQPMYFNEGRGKVVYVELQNRVYNGTLKDSHSYQNA 540

Qy 541 RFGSSIASVDELNODSNDVYVVGAPLEDNHAGAIYIIPHGRGSIILKTPKQITASELATG 600

Db 541 RFGSSIASVDELNODSNDVYVVGAPLEDNHAGAIYIIPHGRGSIILKTPKQITASELATG 600

Qy 601 LOYFGCSIHGOLDLNEGLIDLAVCALGNVILMSRPVVOQINASLHFPSPKINIFHRDCK 660

Db 601 LOYFGCSIHGOLDLNEGLIDLAVCALGNVILMSRPVVOQINASLHFPSPKINIFHRDCK 660

Qy 661 RSGRDATCLAAFLCFTPIFLAPHPQTIVGIRYNATMDERRYTTPRAHLDEGDRFTNRVAV 720

Db 661 RSGRDATCLAAFLCFTPIFLAPHPQTIVGIRYNATMDERRYTTPRAHLDEGDRFTNRVAV 720

Qy 721 LLSGQELCERINPHVLDADYVVPFVSVEYSLEDDPHGPMLDGOWTTLRVSVFPFWNG 780

Db 721 LLSGQELCERINPHVLDADYVVPFVSVEYSLEDDPHGPMLDGOWTTLRVSVFPFWNG 780

Qy 781 CNEDEHCVPLDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLSPDTTVPFIIESTRQVAV 840

Db 781 CNEDEHCVPLDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLSPDTTVPFIIESTRQVAV 840

Qy 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNERRLOKQVCNVSYPFP 900

Db 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNERRLOKQVCNVSYPFP 900

Qy 901 RAKAKVAPRLDSEPSKSIPLJHLEIELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFT 960

Db 901 RAKAKVAPRLDSEPSKSIPLJHLEIELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFT 960

Qy 961 RSSLSHYEVKLSLSERYDGIQPPFSCIPRIQNLGLPPIHGMKMKITIPATRSNRL 1020

Db 961 RSSLSHYEVKLSLSERYDGIQPPFSCIPRIQNLGLPPIHGMKMKITIPATRSNRL 1020

Qy 1021 KLRFPLTDE-ANTSNIWGNSTERYPTVEEDLARAOLAHNSDVVSNICNIRLVPNOE 1079

Db 1021 KLRFPLTDE-ANTSNIWGNSTERYPTVEEDLARAOLAHNSDVVSNICNIRLVPNOE 1080

Qy 1080 INFHLGNLWLSLKALKYKSMKIMVNAALQORQFHSPFFIFREBDDPSRQIVFISKQEDWQ 1139

Db 1081 INFHLGNLWLSLKALKYKSMKIMVNAALQORQFHSPFFIFREBDDPSRQIVFISKQEDWQ 1140

Qy 1140 VPIWIIIVGSTLGGILLALLVLALRLKLGPPRSARRRREBGLDTPKYLE 1188

Db 1141 VPIWIIIVGSTLGGILLALLVLALWKLGGFRSARRRREBGLDTPKYLE 1189

RESULT 5

US-10-262-839-4

/ Sequence 4, Application US/10262839

/ Publication No. US20040038877A1

/ GENERAL INFORMATION:

/ APPLICANT: Alsobrook, John,

/ APPLICANT: Anderson, David W.,

/ APPLICANT: Boldog, Ferenc,

/ APPLICANT: Burgess, Catherine,

/ APPLICANT: Catterton, Elina,

/ APPLICANT: Edinger, Shlomit,

/ APPLICANT: Ellerman, Karen,

/ APPLICANT: Gerlach, Valerie,

/ APPLICANT: Gorman, Linda,

/ APPLICANT: Guo, Xiaojia,

/ APPLICANT: Ji, Weizhen,

/ APPLICANT: Kekoda, Ramesh,

/ APPLICANT: Leach, Martin,

/ APPLICANT: Li, Li,

/ APPLICANT: Miller, Charles,

/ APPLICANT: Patturajan, Meera,

/ APPLICANT: Reiger, Daniel,

/ APPLICANT: Rothenberg, Mark,

/ APPLICANT: Shinkets, Richard,

/ APPLICANT: Smithson, Glennda,

/ APPLICANT: Spytek, Kimberly,

/ APPLICANT: Taupier, Raymond, jr.,

/ APPLICANT: Vernet, Corine,

/ APPLICANT: Voss, Edward,

/ APPLICANT: Zerhusen, Brian,



APPLICANT: Zhong, Mei  
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
FILE REFERENCE: 21402-462A  
CURRENT APPLICATION NUMBER: US/10/262,839  
CURRENT FILING DATE: 2002-10-01  
PRIOR APPLICATION NUMBER: 60/326,483  
PRIOR FILING DATE: 2001-10-02  
PRIOR APPLICATION NUMBER: 60/327,917  
PRIOR FILING DATE: 2001-10-09  
PRIOR APPLICATION NUMBER: 60/328,029  
PRIOR FILING DATE: 2001-10-09  
PRIOR APPLICATION NUMBER: 60/328,056  
PRIOR FILING DATE: 2001-10-09  
PRIOR APPLICATION NUMBER: 60/381,101  
PRIOR FILING DATE: 2002-05-16  
PRIOR APPLICATION NUMBER: 60/371,972  
PRIOR FILING DATE: 2002-04-12  
PRIOR APPLICATION NUMBER: 60/327,342  
PRIOR FILING DATE: 2001-10-05  
PRIOR APPLICATION NUMBER: 60/328,044  
PRIOR FILING DATE: 2001-10-09  
PRIOR APPLICATION NUMBER: 60/328,849  
PRIOR FILING DATE: 2001-10-12  
PRIOR APPLICATION NUMBER: 60/374,738  
PRIOR FILING DATE: 2002-04-23  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 367  
SOFTWARE: CuraSeqList version 0.1  
SEQ ID NO 4  
LENGTH: 1189  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-262-839-4

Query Match 99.4%; Score 6188.5; DB 12; Length 1189;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1184; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
QY 1 MDPLGLVAVVWALSHPGFTDTFNDTRKPRVPGSRPTAFTGYTQQHDSGNKWL VVGA 60  
DB 1 MDPLGLVAVVWALSHPGFTDTFNDTRKPRVPGSRPTAFTGYTQQHDSGNKWL VVGA 60  
QY 61 PLETNGYQGTGVYKCPVHGNCTKLNIGRVTLSNVSRKDNMRGLSLATNPKNDSFLA 120  
DB 61 PLETNGYQGTGVYKCPVHGNCTKLNIGRVTLSNVSRKDNMRGLSLATNPKNDSFLA 120  
QY 121 CSPLWSHCGSSYTTGCMRSVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSNTPVVE 180  
DB 121 CSPLWSHCGSSYTTGCMRSVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSNTPVVE 180  
QY 181 VQHFLINILKXPIYIGQIQGVQYGEDVWVEHFLNDYRSYKOVVEAASHIBQGGTET 240  
DB 181 VQHFLINILKXPIYIGQIQGVQYGEDVWVEHFLNDYRSYKOVVEAASHIBQGGTET 240  
QY 241 RTARGIEFARSAPFQGRKGAKKVMIVITDGHSDSPDLEKVIQOOSRDNVRYAVAVL 300  
DB 241 RTARGIEFARSAPFQGRKGAKKVMIVITDGHSDSPDLEKVIQOOSRDNVRYAVAVL 300  
QY 301 GYNNRGINPTFLNEIKYIASDDPKHFFNVNTDEAALKOIVDALGDRIFSLGKNNKET 360  
DB 301 GYNNRGINPTFLNEIKYIASDDPKHFFNVNTDEAALKOIVDALGDRIFSLGKNNKET 360  
QY 361 SFGLMSQTFSSHVVDGVLGAVGAYDNGAVLKETSAKVIPLESYLYKEPPEELKN 420  
DB 361 SFGLMSQTFSSHVVDGVLGAVGAYDNGAVLKETSAKVIPLESYLYKEPPEELKN 420  
QY 421 HCAYLGYYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTWNNRSLTIHQAMRGQOIGSYF 480  
DB 421 HCAYLGYYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTWNNRSLTIHQAMRGQOIGSYF 480  
QY 481 GSEITSVIDGGVTDVLLVAGAPMYFNEGRGKVVYVELQNRPNVYNTLKDSHSYONA 540  
DB 481 GSEITSVIDGGVTDVLLVAGAPMYFNEGRGKVVYVELQNRPNVYNTLKDSHSYONA 540

QY 541 RFSSIASVRLDNLQDSYNDVWVGAPLEDNAGAIYIFHGRGSIILTKPKORITASELATG 600  
DB 541 RFSSIASVRLDNLQDSYNDVWVGAPLEDNAGAIYIFHGRGSIILTKPKORITASELATG 600  
QY 601 LOYFGGSIHGOLDLNEGLDLAVGALGNVILWSRPVQINASLHFEPPSKINIFHRDCK 660  
DB 601 LOYFGGSIHGOLDLNEGLDLAVGALGNVILWSRPVQINASLHFEPPSKINIFHRDCK 660  
QY 661 RSGRDATCLAAFLCPTPIFLAPHQTTTIGIRNATWDERRYPRAHLDDGGDRFTTRAV 720  
DB 661 RSGRDATCLAAFLCPTPIFLAPHQTTTIGIRNATWDERRYPRAHLDDGGDRFTTRAV 720  
QY 721 LSSGQELCSRIHFHVLDTADYVKPTFSVEYSLEDHDPMLDDGPTTLRVSPVFWNG 780  
DB 721 LSSGQELCSRIHFHVLDTADYVKPTFSVEYSLEDHDPMLDDGPTTLRVSPVFWNG 780  
QY 781 CNEDEHCVPLDLVLDARSDDLPTAMEYCORVLRKPAQDCSAYTSLPDTTVFIIESTRQVAV 840  
DB 781 CNEDEHCVPLDLVLDARSDDLPTAMEYCORVLRKPAQDCSAYTSLPDTTVFIIESTRQVAV 840  
QY 841 EATLENREGNAYSTVLNISOSANLQFASLIQKEDSDGSIIECVNEERLQKQVCNVSYPFF 900  
DB 841 EATLENREGNAYSTVLNISOSANLQFASLIQKEDSDGSIIECVNEERLQKQVCNVSYPFF 900  
QY 901 RAKAKVAFRLDSEFSKSIPLHLEIHLAAGSDSNERDSTKEDNVAPLRFHLYEADVLFT 960  
DB 901 RAKAKVAFRLDSEFSKSIPLHLEIHLAAGSDSNERDSTKEDNVAPLRFHLYEADVLFT 960  
QY 961 RSSLSHYEVKLNSSLERYDGIIPPESCIFRIQNLGLFPIHGMWMMKITIPIATSGNRL 1020  
DB 961 RSSLSHYEVKLNSSLERYDGIIPPESCIFRIQNLGLFPIHGMWMMKITIPIATSGNRL 1020  
QY 1021 KLRFDLTDE-ANTSCNINWGNSTBYRPTFVEEDLRRAPQLNHSNSDVVSINCNIRLVPNQ 1079  
DB 1021 KLRFDLTDE-ANTSCNINWGNSTBYRPTFVEEDLRRAPQLNHSNSDVVSINCNIRLVPNQ 1079  
QY 1080 INFHLLGNLWLSLKALKYKSMKIMWNAALQROFHSPIFREEDPSQIIEFISKOEDWQ 1139  
DB 1080 INFHLLGNLWLSLKALKYKSMKIMWNAALQROFHSPIFREEDPSQIIEFISKOEDWQ 1139  
QY 1140 VPIWIIVGSITLGGILLALLVALLRLKGFPRFARRRREPGLDPTPKVLE 1188  
DB 1140 VPIWIIVGSITLGGILLALLVALLRLKGFPRFARRRREPGLDPTPKVLE 1188

RESULT 6  
US-10-262-839-6  
Sequence 6, Application US/10262839  
Publication No. US20040038877A1  
GENERAL INFORMATION:  
APPLICANT: Alsobrook, John,  
APPLICANT: Anderson, David W.,  
APPLICANT: Boldog, Ferenc,  
APPLICANT: Burgess, Catherine,  
APPLICANT: Catterton, Elina,  
APPLICANT: Edinger, Shlomit,  
APPLICANT: Ellerman, Karen,  
APPLICANT: Gerlach, Valerie,  
APPLICANT: Guo, Xiaojia,  
APPLICANT: Ji, Weizhen,  
APPLICANT: Kekuda, Ramesh,  
APPLICANT: Leach, Martin,  
APPLICANT: Li, Li, Charles,  
APPLICANT: Miller, Charles,  
APPLICANT: Patturajan, Meera,  
APPLICANT: Reiger, Daniel,  
APPLICANT: Rothenberg, Mark,  
APPLICANT: Shimkets, Richard,  
APPLICANT: Smithson, Glennda,  
APPLICANT: Spytek, Kimberly,  
APPLICANT: Taupier, Raymond, Jr.,

```

; APPLICANT: Vernet, Corine,
; APPLICANT: Voss, Edward,
; APPLICANT: Zertusen, Brian,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-462A
; CURRENT APPLICATION NUMBER: US/10/262,839
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,101
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/371,972
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/327,342
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/328,044
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/374,738
; PRIOR FILING DATE: 2002-04-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: CuraSequid version 0.1
; SEQ ID NO 6
; LENGTH: 1120
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-262-839-6

Query Match      92.9%; Score 5780; DB 12; Length 1120;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1115; Conservative 1; Mismatches 3; Indels 70; Gaps 2;

Qy      1 MDLPRGLVVAWALSINLPGFTDFTNMDTRKPRVIGSRFAFFGYVQQHDSGNKMLVUGA 60
Db      1 MDLPRGLVVAWALSINLPGFTDFTNMDTRKPRVIGSRFAFFGYVQQHDSGNKMLVUGA 60

Qy      61 PLETNGYQKTDVYKCPVHGNCTKLALGRVTLNVSERKDNMRLGLSLATNPKNDSFLA 120
Db      61 PLETNGYQKTDVYKCPVHGNCTKLNLG----- 89

Qy      121 CSLPLWSHECGSSYTTGMCSSRVNSNFRPSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
Db      90 -----CQTYMDIVIVLDGNSIYPWVE 111

Qy      181 VQHFELINLKEFYIGPGQIQGVGVQYGEDVHVHEFLNDYRSKDVWEAASHIEQGGTET 240
Db      112 VQHFELINLKEFYIGPGQIQGVGVQYGEDVHVHEFLNDYRSKDVWEAASHIEQGGTET 171

Qy      241 RTAFGIEPARSEAFQKGRKAKKVMIVITDGSCHSDPDLEKVIQQSERDNVTRYAVAVL 300
Db      172 RTAFGIEPARSEAFQKGRKAKKVMIVITDGSCHSDPDLEKVIQQSERDNVTRYAVAVL 231

Qy      301 GYNRRGINPPEFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSLGCTNKNET 360
Db      232 GYNRRGINPPEFLNEIKYIASDDDDKHFFNVTDAAALKDIDVALGDRIFSLGCTNKNET 291

Qy      361 SFGLMSGTGFSSHHVEDGVLLGAVGYDMNGAVLKETSAGKVIPLRESYKGFPEELKN 420
Db      292 SFGLMSGTGFSSHHVEDGVLLGAVGYDMNGAVLKETSAGKVIPLRESYKGFPEELKN 351

Qy      421 HGAYLGVTVTSVSSRGRVTVAGAPRPNHTGKVLFTMNNRSLTTHQAMRQQIGSYF 480
Db      352 HGAYLGVTVTSVSSRGRVTVAGAPRPNHTGKVLFTMNNRSLTTHQAMRQQIGSYF 411

```

RESULT 7

```

US-09-984-130-43
; Sequence 43, Application US/09984130
; Publication No. US20030055231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489B2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 1034

```

```

Qy      481 GSEITSDVIDGQVTDVLLVGAPEYMFNGBRGKVVYVELRONRFFVYNGTLKDSHSYQNA 540
Db      412 GSEITSDVIDGQVTDVLLVGAPEYMFNGBRGKVVYVELRONRFFVYNGTLKDSHSYQNA 471

Qy      541 RFGSSIASVRDLNQDSYNDVVVGAFLDNDHAGAIYIFHGFGRGSIILKTPQRITASELATG 600
Db      472 RFGSSIASVRDLNQDSYNDVVVGAFLDNDHAGAIYIFHGFGRGSIILKTPQRITASELATG 531

Qy      601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILSRPVVQINASHLFBSKINIPIHRDCK 660
Db      532 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILSRPVVQINASHLFBSKINIPIHRDCK 591

Qy      661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIIRYNATMDERRYPRAHLDEGGDRFNRAV 720
Db      592 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIIRYNATMDERRYPRAHLDEGGDRFNRAV 651

Qy      721 LLSGQELCERINFHVLDYADYVYKPVTFSEVLSLEDDPHGPMDDGWPTTLRVSVFFWNG 780
Db      652 LLSGQELCERINFHVLDYADYVYKPVTFSEVLSLEDDPHGPMDDGWPTTLRVSVFFWNG 711

Qy      781 CNEDEHCVPLDLVLDARSOLPTAMEYCORVLRKPAODCSAYTILSFDTTVEIIESTRQAV 840
Db      712 CNEDEHCVPLDLVLDARSOLPTAMEYCORVLRKPAODCSAYTILSFDTTVEIIESTRQAV 771

Qy      841 EATLENRGENAYSTVLNISOANLQFASLIQKEDSDGSIQCVNEERLQKQVNSYPPF 900
Db      772 EATLENRGENAYSTVLNISOANLQFASLIQKEDSDGSIQCVNEERLQKQVNSYPPF 831

Qy      901 RAKAKVAPRLDSEPSKSIPLHLEIETLAGSUSNERDSTKEONVAPLRHLXVADVLFT 960
Db      832 RAKAKVAPRLDSEPSKSIPLHLEIETLAGSUSNERDSTKEONVAPLRHLXVADVLFT 891

Qy      961 RSSLSHSHVYKLNLSLERYDGIQPPSCIFRIQNLGLFPIHGMMKTIPIATRSNRL 1020
Db      892 RSSLSHSHVYKLNLSLERYDGIQPPSCIFRIQNLGLFPIHGMMKTIPIATRSNRL 951

Qy      1021 KLKDLFLTDE-ANTSCNINWGNSTERYPTVVEEDLRAPQLNHSNDVVSINCMRLVPOE 1079
Db      952 KLKDLFLTDEVANTSCNINWGNSTERYPTVVEEDLRAPQLNHSNDVVSINCMRLVPOE 1011

Qy      1080 INFHLIGNLWLSLAKALKYKSMKIWNVAALQRFSPPIFRBEDPSRQIEFEISKQEDWQ 1139
Db      1012 INFHLIGNLWLSLAKALKYKSMKIWNVAALQRFSPPIFRBEDPSRQIEFEISKQEDWQ 1071

Qy      1140 VPIWIIIVGTSTIGLLILLALLVLALRKLGFRRARRRREPCLDPTPKVLE 1188
Db      1072 VPIWIIIVGTSTIGLLILLALLVLALRKLGFRRARRRREPCLDPTPKVLE 1120

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-43

Query Match      86.5%; Score 5383; DB 10; Length 1034;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1026; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDLPGLVAVWALSHPGFTDTFNMDTRKPRVPGSRTPAFGTVVQOHDISGNKWLVVGA 60
DB 1 MDLPGLVAVWALSHPGFTDTFNMDTRKPRVPGSRTPAFGTVVQOHDISGNKWLVVGA 60
QY 61 PLENGYOKTGVYKCPVHGNCTKLNLRVTLNVSRKONMRGLSLATNPKDNSFLA 120
DB 61 PLENGYOKTGVYKCPVHGNCTKLNLRVTLNVSRKONMRGLSLATNPKDNSFLA 120
QY 121 CSPLSHSCGSSYYTTGMCSSRVNSFRFSKTVAPALQRCQTYMDIVVLGDSNSIYPWVE 180
DB 121 CSPLSHSCGSSYYTTGMCSSRVNSFRFSKTVAPALQRCQTYMDIVVLGDSNSIYPWVE 180
QY 181 VOHFLINILKKFYIGPGQIOGVVQYGEDVHGFHNDYRSKDVVEAASHIEBORGTTET 240
DB 181 VOHFLINILKKFYIGPGQIOGVVQYGEDVHGFHNDYRSKDVVEAASHIEBORGTTET 240
QY 181 VOHFLINILKKFYIGPGQIOGVVQYGEDVHGFHNDYRSKDVVEAASHIEBORGTTET 240
DB 181 VOHFLINILKKFYIGPGQIOGVVQYGEDVHGFHNDYRSKDVVEAASHIEBORGTTET 240
QY 241 RTAFGIEFARSEAFQKGRGKAKVMIVITDGESHDSPLDKVIOQSERDNVTRYAVAVL 300
DB 241 RTAFGIEFARSEAFQKGRGKAKVMIVITDGESHDSPLDKVIOQSERDNVTRYAVAVL 300
QY 301 GYNRRGINPETFLNEIKYIASDDPKHFFNTDEAALKDIDVALGDRIFSLGNTKNET 360
DB 301 GYNRRGINPETFLNEIKYIASDDPKHFFNTDEAALKDIDVALGDRIFSLGNTKNET 360
QY 361 SFGLMSOTGFSHVVEDGVLGAVGAYDNGAVLAKTSAGKVIPLRESYLKSPPEELKN 420
DB 361 SFGLMSOTGFSHVVEDGVLGAVGAYDNGAVLAKTSAGKVIPLRESYLKSPPEELKN 420
QY 421 HGYALGYTVTSVSSRQGRVYVAGAPFNHTGKVIPLTMNRRSLTIHQMRGQIGSYF 480
DB 421 HGYALGYTVTSVSSRQGRVYVAGAPFNHTGKVIPLTMNRRSLTIHQMRGQIGSYF 480
QY 481 RAKAVAFRLDSERKSFLEHLETELAAGSDSDERDSTKEDNVAPLRFHLKYEADVLFT 960
DB 481 RAKAVAFRLDSERKSFLEHLETELAAGSDSDERDSTKEDNVAPLRFHLKYEADVLFT 960
QY 961 RSSLSHYEVKLNSSLERYDGGPPPCIFRIQNLGLFPFINGIMMKITPIATRSNRL 1020

; Sequence 43, Application US/09836353A
; Publication No. US20030129685A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P1
; CURRENT APPLICATION NUMBER: US/09/836.353A
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-353A-43

Query Match      86.5%; Score 5383; DB 10; Length 1034;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1026; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDLPGLVAVWALSHPGFTDTFNMDTRKPRVPGSRTPAFGTVVQOHDISGNKWLVVGA 60
DB 1 MDLPGLVAVWALSHPGFTDTFNMDTRKPRVPGSRTPAFGTVVQOHDISGNKWLVVGA 60
QY 61 PLENGYOKTGVYKCPVHGNCTKLNLRVTLNVSRKONMRGLSLATNPKDNSFLA 120
DB 61 PLENGYOKTGVYKCPVHGNCTKLNLRVTLNVSRKONMRGLSLATNPKDNSFLA 120
QY 121 CSPLSHSCGSSYYTTGMCSSRVNSFRFSKTVAPALQRCQTYMDIVVLGDSNSIYPWVE 180
DB 121 CSPLSHSCGSSYYTTGMCSSRVNSFRFSKTVAPALQRCQTYMDIVVLGDSNSIYPWVE 180
QY 181 VOHFLINILKKFYIGPGQIOGVVQYGEDVHGFHNDYRSKDVVEAASHIEBORGTTET 240
DB 181 VOHFLINILKKFYIGPGQIOGVVQYGEDVHGFHNDYRSKDVVEAASHIEBORGTTET 240
QY 181 VOHFLINILKKFYIGPGQIOGVVQYGEDVHGFHNDYRSKDVVEAASHIEBORGTTET 240
DB 181 VOHFLINILKKFYIGPGQIOGVVQYGEDVHGFHNDYRSKDVVEAASHIEBORGTTET 240
QY 241 RTAFGIEFARSEAFQKGRGKAKVMIVITDGESHDSPLDKVIOQSERDNVTRYAVAVL 300
DB 241 RTAFGIEFARSEAFQKGRGKAKVMIVITDGESHDSPLDKVIOQSERDNVTRYAVAVL 300
QY 301 GYNRRGINPETFLNEIKYIASDDPKHFFNTDEAALKDIDVALGDRIFSLGNTKNET 360
DB 301 GYNRRGINPETFLNEIKYIASDDPKHFFNTDEAALKDIDVALGDRIFSLGNTKNET 360
QY 361 SFGLMSOTGFSHVVEDGVLGAVGAYDNGAVLAKTSAGKVIPLRESYLKSPPEELKN 420
DB 361 SFGLMSOTGFSHVVEDGVLGAVGAYDNGAVLAKTSAGKVIPLRESYLKSPPEELKN 420
QY 421 HGYALGYTVTSVSSRQGRVYVAGAPFNHTGKVIPLTMNRRSLTIHQMRGQIGSYF 480
DB 421 HGYALGYTVTSVSSRQGRVYVAGAPFNHTGKVIPLTMNRRSLTIHQMRGQIGSYF 480
QY 481 RAKAVAFRLDSERKSFLEHLETELAAGSDSDERDSTKEDNVAPLRFHLKYEADVLFT 960
DB 481 RAKAVAFRLDSERKSFLEHLETELAAGSDSDERDSTKEDNVAPLRFHLKYEADVLFT 960
QY 961 RSSLSHYEVKLNSSLERYDGGPPPCIFRIQNLGLFPFINGIMMKITPIATRSNRL 1020
```

```
Db 541 RFSSIASVRLDNDKSDVNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKQITASELATG 600
Qy 601 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPVQVNASLHFEPSKINIFHRDCK 660
Db 601 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPVQVNASLHFEPSKINIFHRDCK 660
Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGIRYNATMDERYTTPRAHLDGGDRTNRAV 720
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGIRYNATMDERYTTPRAHLDGGDRTNRAV 720
Qy 721 LLSGQGLCERINPHVLDATYVYKPVTFSEYSELEDPDHGPMLDDGMPPTLRVSPFWNG 780
Db 721 LLSGQGLCERINPHVLDATYVYKPVTFSEYSELEDPDHGPMLDDGMPPTLRVSPFWNG 780
Qy 781 CNEDEHCVPLVDARSDLPTAMEYCORVLKPAQCCSAVTLSDFTTPTVFIESTORVAV 840
Db 781 CNEDEHCVPLVDARSDLPTAMEYCORVLKPAQCCSAVTLSDFTTPTVFIESTORVAV 840
Qy 841 EATLENKGENAYSTVLNISQSANLQFASLIQKSDSDGSIIECVNEERLQKQVCNVSYPFF 900
Db 841 EATLENKGENAYSTVLNISQSANLQFASLIQKSDSDGSIIECVNEERLQKQVCNVSYPFF 900
Qy 901 RAKAKVAPRLDSPEKSI FLHLEIELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFT 960
Db 901 RAKAKVAPRLDSPEKSI FLHLEIELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFT 960
Qy 961 RSSLSHYEVKLANSRLERYDGIQPPFSCIERIQLGLFPIHGMKITTPIATRSNRL 1020
Db 961 RSSLSHYEVKLANSRLERYDGIQPPFSCIERIQLGLFPIHGMKITTPIATRSNRL 1020
Qy 1021 KLKRDFLTDE 1029
Db 1021 KLKRDFLTDE 1029
```

## RESULT 9

```
US-09-764-870-313
; Sequence 313, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 313
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-313
```

```
Query Match 58.2%; Score 3621; DB 9; Length 707;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 689; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 MDLPRLVVAWALSLLWPGFTDTFNMDTRKPRVTPGSRTPAFPGYTVQOHDISGNKWLWVGA 60
Db 17 MDLPRLVVAWALSLLWPGFTDTFNMDTRKPRVTPGSRTPAFPGYTVQOHDISGNKWLWVGA 76
Qy 61 PLETNGYOKTGDVYKCPVIHGNCTKLNLRVTLSNVSEKDNKRLGLSLATNPKNSFLA 120
Db 77 PLETNGYOKTGDVYKCPVIHGNCTKLNLRVTLSNVSEKDNKRLGLSLATNPKNSFLA 136
Qy 121 CSPLMSHSCGSSYYTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNLYPWVE 180
Db 137 CSPLMSHSCGSSYYTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNLYPWVE 196
Qy 181 VQHFLINILKKFYIGPGQIQGVQGEDVHVEFLNDYRSVKDVVEAASHIEQRGGTET 240
```

```
Db 197 VQHFLINILKKFYIGPGQIQGVQGEDVHVEFLNDYRSVKDVVEAASHIEQRGGTET 256
Qy 241 RTAFGIEFARSAFQKGRKGAKTVMIVITDGESHSDPLEKVIQQSERDNTVRYAVVL 300
Db 257 RTAFGIEFARSAFQKGRKGAKTVMIVITDGESHSDPLEKVIQQSERDNTVRYAVVL 316
Qy 301 GYNNRRGINPETFLNEIKYIASDDPDKHFFNVVTDAAALKDIDVALGDRIFSLGEGTNKNET 360
Db 317 GYNNRRGINPETFLNEIKYIASDDPDKHFFNVVTDAAALKDIDVALGDRIFSLGEGTNKNET 376
Qy 361 SFGLEMSQTGSSHHVVDGVLGAVGYDMNGAVLKETISAGKVIPLRSYLYKEFPEELKN 420
Db 377 SFGLEMSQTGSSHHVVDGVLGAVGYDMNGAVLKETISAGKVIPLRSYLYKEFPEELKN 436
Qy 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTKGVILFTMNNRSJTIHQAMGQQIGSYF 480
Db 437 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTKGVILFTMNNRSJTIHQAMGQQIGSYF 496
Qy 481 GSEITSDIDGVDVLLVGAPMYFNEGRERKGVYVYVELRQNRVYFNGTLKDSHSYQNA 540
Db 497 GSEITSDIDGVDVLLVGAPMYFNEGRERKGVYVYVELRQNRVYFNGTLKDSHSYQNA 556
Qy 541 RFGSSIASVRDLNDSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKQITASELATG 600
Db 557 RFGSSIASVRDLNDSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKQITASELATG 616
Qy 601 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPVQVNASLHFEPSKINIFHRDCK 660
Db 617 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPVQVNASLHFEPSKINIFHRDCK 676
Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGI 691
Db 677 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGI 707
```

## RESULT 10

```
US-10-125-540-313
; Sequence 313, Application US/10125540
; Publication No. US20030059875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214C1
; CURRENT APPLICATION NUMBER: US/10/125,540
; CURRENT FILING DATE: 2002-04-19
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 313
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-125-540-313
```

```
Query Match 58.2%; Score 3621; DB 14; Length 707;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 689; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 MDLPRLVVAWALSLLWPGFTDTFNMDTRKPRVTPGSRTPAFPGYTVQOHDISGNKWLWVGA 60
Db 17 MDLPRLVVAWALSLLWPGFTDTFNMDTRKPRVTPGSRTPAFPGYTVQOHDISGNKWLWVGA 76
Qy 61 PLETNGYOKTGDVYKCPVIHGNCTKLNLRVTLSNVSEKDNKRLGLSLATNPKNSFLA 120
Db 77 PLETNGYOKTGDVYKCPVIHGNCTKLNLRVTLSNVSEKDNKRLGLSLATNPKNSFLA 136
Qy 121 CSPLMSHSCGSSYYTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNLYPWVE 180
Db 137 CSPLMSHSCGSSYYTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNLYPWVE 196
Qy 181 VQHFLINILKKFYIGPGQIQGVQGEDVHVEFLNDYRSVKDVVEAASHIEQRGGTET 240
Db 197 VQHFLINILKKFYIGPGQIQGVQGEDVHVEFLNDYRSVKDVVEAASHIEQRGGTET 256
```

QY 241 RTAFGIEPARSEAFQKGRGKAKVMIVITDGHSDSPDLKVIQOSREDNVTYAVAVL 300  
DB 257 RTAFGIEPARSEAFQKGRGKAKVMIVITDGHSDSPDLKVIQOSREDNVTYAVAVL 316  
QY 301 GYNRRGINPETFLNEIKVIASDPDDKHFNFVNTDEAALKDVIDALGDRIPSELTGNET 360  
DB 317 GYNRRGINPETFLNEIKVIASDPDDKHFNFVNTDEAALKDVIDALGDRIPSELTGNET 376  
QY 361 SFGLEMSQGFSSHVVYDGVLLGAVGAYDNGAVLKETSAGKVIPLRBSYLKPEPEELKN 420  
DB 377 SFGLEMSQGFSSHVVYDGVLLGAVGAYDNGAVLKETSAGKVIPLRBSYLKPEPEELKN 436  
QY 421 HGAYLGTYTTSVSSRQGRVYVAGAPRFNHTGKVLFTMHNNRSLTIHQAMEGQOIGSYF 480  
DB 437 HGAYLGTYTTSVSSRQGRVYVAGAPRFNHTGKVLFTMHNNRSLTIHQAMEGQOIGSYF 496  
QY 481 GSEITSDVDGQVTDVLLGAPMVFNEGRECKVYVYELQRNRFYNGTLDKDSHYQNA 540  
DB 497 GSEITSDVDGQVTDVLLGAPMVFNEGRECKVYVYELQRNRFYNGTLDKDSHYQNA 556  
QY 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFSGSILKTPKQRITASELATG 600  
DB 557 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFSGSILKTPKQRITASELATG 616  
QY 601 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILWSPVQVQINASLHFSPSKINIPIHRDCK 660  
DB 617 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILWSPVQVQINASLHFSPSKINIPIHRDCK 676  
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTNGI 691  
DB 677 RSGRDATCLAAFLCFTPIFLAPHFQTTTNGI 707

## RESULT 11

US-09-866-050A-624

; Sequence 624, Application US/09866050A

; Publication No. US2003004071A1

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Murison, James G.

; APPLICANT: Kumble, Krishanand D.

; TITLE OF INVENTION: Compositions Isolated From Skin Cells

; TITLE OF INVENTION: and Methods for Their Use

; FILE REFERENCE: 11000.1011c4U

; CURRENT FILING DATE: 2001-05-24

; NUMBER OF SEQ ID NOS: 725

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 624

; LENGTH: 688

; TYPE: PRT

; ORGANISM: Rat

US-09-866-050A-624

Query Match 51.1%; Score 3182; DB 10; Length 688;

Best Local Similarity 86.5%; Pred. No. 2.6e-284;

Matches 595; Conservative 46; Mismatches 47; Indels 0; Gaps 0;

QY 501 GAPMYFNEGERGKVVYVYELQRNRFYNGTLDKDSHYQNAFSGSSIASVRDLNQDSYNDV 560

DB 1 GAPMYFNEGERGKVVYVYELQRNRFYNGTLDKDSHYQNAFSGSSIASVRDLNQDSYNDV 60

QY 561 VVGAPLEDNHAGAIYIFHGFSGSILKTPKQRITASELATGLOYPGCSIHGQDLNEDGLI 620

DB 61 VVGAPLEDNHAGAIYIFHGFSGSILKTPKQRITASELATGLOYPGCSIHGQDLNEDGLV 120

QY 621 DLAVGALGNVILWSPVQVQINASLHFSPSKINIPIHRDCKSGRDATCLAAFLCFTPIFL 680

DB 121 DLAVGALGNVILWSPVQVQINASLHFSPSKINIPIHRDCKSGRDATCLAAFLCFTPIFL 180

QY 681 APHECTTIVGIRYNATMDREYVTPRAHLDEGDBFTNRAVLLSSQELCERINFHVLDTA 740  
DB 181 APHECTTIVGIRYNATMDREYVTPRAHLDEGDBFTNRAVLLSSQELCERINFHVLDTA 240  
QY 741 DYVXPVTFVSYSLEDPDHGPMLDDGMPTTLRVSVPPFWMGNCNDEHCHVDPDLVLDARSDFP 800  
DB 241 DYVXPVTFVSYSLEDPDHGPMLDDGMPTTLRVSVPPFWMGNCNDEHCHVDPDLVLDARSDFP 300  
QY 801 TAMEYQORVLRKPAQDCSAVTLSPDVTTFVFIESTRORVAVATLENRCENAYSTVLTNSQ 860  
DB 301 TAMEYQORVLRKPAQDCSSYTLSPDVTTFVFIESTRORVAVATLENRCENAYSTVLTNSQ 360  
QY 861 SANLQFASLIQKESBDSGSEICVNEERLQKQCNVSVPPFRAKAKVAFRLDSSEFSKISFL 920  
DB 361 SENLQFASLIQKDSBDSGSEICVNEERLQKQCNVSVPPFRAKAKVAFRLDSSEFSKISFL 420  
QY 921 HHLXIELAAGSDNSHERDSTKEDNVAPLRFHLKYEADVLFTRSSSLSHYEVKLNSSLEPYD 980  
DB 421 HHLXIELAAGSDNSHERDSTKEDNVAPLRFHLKYEADVLFTRSSSLSHYEVKLNSSLEPYD 480  
QY 981 GIGPPFSCIFRQNLGLFPPIHGMWKITPIATSGNRLLKLRDPLTDANTSCNMGNS 1040  
DB 481 GIGPPFSCIFRQNLGLFPPIHGMWKITPIATSGNRLLKLRDPLTDANTSCNMGNS 540  
QY 1041 TEYRPTVVEEDLRAPQNLHNSDVVVSINCNIRLVNQBENHLLGNLWLSLKALKYKS 1100  
DB 541 TEYRPTVVEEDLRAPQNLHNSDVVVSINCNIRLVNQBENHLLGNLWLSLKALKYKS 600  
QY 1101 MKIMVNAALQRFHSPRIFREEDPSRQIEFEISKOEDQVPIWIIIVGSTLGLLALLLV 1160  
DB 601 MKIMVNAALQRFHSPRIFREEDPSRQIEFEISKOEDQVPIWIIIVGSTLGLLALLLV 660  
QY 1161 LALRKLGFRRSARRRREPGLDPTPKVLE 1188  
DB 661 LALRKLGFRRSARRRREPGLDPTPKVLE 688

## RESULT 12

US-09-866-050A-501

; Sequence 501, Application US/09866050A

; Publication No. US2003004071A1

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Murison, James G.

; APPLICANT: Kumble, Krishanand D.

; TITLE OF INVENTION: Compositions Isolated From Skin Cells

; TITLE OF INVENTION: and Methods for Their Use

; FILE REFERENCE: 11000.1011c4U

; CURRENT APPLICATION NUMBER: US/09/866,050A

; CURRENT FILING DATE: 2001-05-24

; NUMBER OF SEQ ID NOS: 725

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 501

; LENGTH: 696

; TYPE: PRT

; ORGANISM: Rat

US-09-866-050A-501

Query Match

Best Local Similarity

Matches 584; Conservative

50.0%; Score 3113; DB 10; Length 696;

85.3%; Pred. No. 6.4e-278;

45; Mismatches 56; Indels 0; Gaps 0;

QY 501 GAPMYFNEGERGKVVYVYELQRNRFYNGTLDKDSHYQNAFSGSSIASVRDLNQDSYNDV 560

DB 1 GAPMYFNEGERGKVVYVYELQRNRFYNGTLDKDSHYQNAFSGSSIASVRDLNQDSYNDV 60

QY 561 VVGAPLEDNHAGAIYIFHGFSGSILKTPKQRITASELATGLOYPGCSIHGQDLNEDGLI 620

DB 61 VVGAPLEDNHAGAIYIFHGFSGSILKTPKQRITASELATGLOYPGCSIHGQDLNEDGLV 120

QY 621 DLAVGALGNVILWSBPVOINASLHFEPSKINI FHRDCKRSGDATCLAAFLCPTPIEL 680  
DB 121 DLAVGALGNVILWAPVVOINASLHFEPSKINI FHRDCKRGRDATCLAAFLCPTPIEL 180  
QY 681 APHFQTTVIGIRYNATMDERRYTPRAHLDGGDRFTNRAVLLSSGQELCERINFHVLDTA 740  
DB 181 APHFQTTVIGIRYNATMDERRYTPRAHLDGGDRFTNRAVLLSSGQELCERINFHVLDTA 240  
QY 741 DYKVPYTFESVEYSLSDPDHGMDDGWPPTLRVSVFPWNGCNEDEKCVDPVLVDARSDEL 800  
DB 241 DYKVPYTFESVEYSLSDPDHGMDDGWPPTLRVSVFPWNGCNEDEKCVDPVLVDARSDEL 300  
QY 801 TAMEYQCVLRKEPAQDCSAYTSLFDTTVPFIIESTRORVAVEATLENRGENAYSTVLNLSQ 860  
DB 301 TAMEYQCVLRKEPAQDCSAYTSLFDTTVPFIIESTRORVAVEATLENRGENAYSTVLNLSQ 360  
QY 861 SANLOFASLIQKSDSGSTECVNEERLQOVKNVSVFPFRKAKAVAFRLDSEFSKSIPL 920  
DB 361 SANLOFASLIQKSDSGSTECVNEERLQOVKNVSVFPFRKAKAVAFRLDSEFSKSIPL 420  
QY 921 HLEIEILAAGSDSNERDSTYKEDNVAPLRFHLKYEADVLFTRSSLSHVEVKNLSLERYD 980  
DB 421 HLEIEILAAGSDSNERDSTYKEDNVAPLRFHLKYEADVLFTRSSLSHVEVKNLSLERYD 480  
QY 981 GIGPPSCIPRIQNLGLFPIHGMKITTPIATRSNRLKLRDPLTDEANTSCNIWNS 1040  
DB 481 GIGPPSCIPRIQNLGLFPIHGMKITTPIATRSNRLKLRDPLTDEANTSCNIWNS 540  
QY 1041 TEYRTPVEEDLRAPOLHNSDDVYSINCNRLVPOQINPHLLNMLWLSLKALKYKS 1100  
DB 541 TEYRTPVEEDLRAPOLHNSDDVYSINCNRLVPOQINPHLLNMLWLSLKALKYKS 600  
QY 1101 MKITMVAALQROPHSPFIREDPSQRIEFSKQEDWQVPIWIIVGSTLGLLLALLV 1160  
DB 601 MKITMVAALQROPHSPFIREDPSQRIEFSKQEDWQVPIWIIVGSTLGLLLALLV 660  
QY 1161 LALRKLGFPSARRRRRGIDPTPK 1185  
DB 661 LALRKLGFPSARRRRRGIDPTPK 685

## RESULT 13

US-09-866-050A-500  
; Sequence 500, Application US/09866050A  
; Publication No. US20030040471A1

## GENERAL INFORMATION:

; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; TITLE OF INVENTION: and Methods for Their Use  
; FILE REFERENCE: 11000.1011C4U  
; CURRENT APPLICATION NUMBER: US/09/866,050A  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 500  
; LENGTH: 545  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-866-050A-500

Query Match 40.5%; Score 2518; DB 10; Length 545;  
Best Local Similarity 86.2%; Pred. No. 4, 2e-223;  
Matches 470; Conservative 35; Mismatches 40; Indels 0; Gaps 0;

QY 644 SLHFEPKINIFHRDCKRGRDATCLAAFLCPTPIELAPHFQTTVIGIRYNATMDERRY 703  
DB 1 SLHFEPKINIFHRDCKRGRDATCLAAFLCPTPIELAPHFQTTVIGIRYNATMDERRY 60

QY 704 PRAHLDGGDRFTNRAVLLSSGQELCERINFHVLDTA DYKVPYTFESVEYSLSDPDHGM 763  
DB 61 PRAHLDGGDRFTNRAVLLSSGQELCERINFHVLDTA DYKVPYTFESVEYSLSDPDHGM 120  
QY 764 DDGWPPTLRVSVFPWNGCNEDEKCVDPVLVDARSDELPTAMEYQCVLRKEPAQDCSAYTSL 823  
DB 121 DDGWPPTLRVSVFPWNGCNEDEKCVDPVLVDARSDELPTAMEYQCVLRKEPAQDCSAYTSL 180  
QY 824 DDGWPPTLRVSVFPWNGCNEDEKCVDPVLVDARSDELPTAMEYQCVLRKEPAQDCSAYTSL 883  
DB 181 DDGWPPTLRVSVFPWNGCNEDEKCVDPVLVDARSDELPTAMEYQCVLRKEPAQDCSAYTSL 240  
QY 884 EERLQOVKNVSVFPFRKAKAVAFRLDSEFSKSIPLHLEIEILAAGSDSNERDSTYKEDN 943  
DB 241 EERLQOVKNVSVFPFRKAKAVAFRLDSEFSKSIPLHLEIEILAAGSDSNERDSTYKEDN 300  
QY 944 VAPLRFHLKYEADVLFTRSSLSHVEVKNLSLERYD GIGPPSCIPRIQNLGLFPIHGM 1003  
DB 301 VAPLRFHLKYEADVLFTRSSLSHVEVKNLSLERYD GIGPPSCIPRIQNLGLFPIHGM 360  
QY 1004 MKITMVAALQROPHSPFIREDPSQRIEFSKQEDWQVPIWIIVGSTLGLLLALLV 1183  
DB 361 MKITMVAALQROPHSPFIREDPSQRIEFSKQEDWQVPIWIIVGSTLGLLLALLV 420  
QY 1183 PSRQVTFEISKQEDWQVPIWIIVGSTLGLLLALLV LALRKLGFPSARRRRRGIDPTPK 1183  
DB 481 PSRQVTFEISKQEDWQVPIWIIVGSTLGLLLALLV LALRKLGFPSARRRRRGIDPTPK 540  
QY 1184 PKVLE 1188  
DB 541 PKVLE 545

## RESULT 14

US-10-108-260A-3386  
; Sequence 3386, Application US/10108260A  
; Publication No. US20040005560A1

## GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1el full length CDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 3386  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-3386

Query Match 35.3%; Score 2200; DB 15; Length 437;  
Best Local Similarity 99.5%; Pred. No. 7e-194;  
Matches 417; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDLPRGLVAVWALSINPFGTDTFNMOTRKPRVIPGSRTPAFPGYTVQOHDISGNKWL VYGA 60  
DB 1 MDLPRGLVAVWALSINPFGTDTFNMOTRKPRVIPGSRTPAFPGYTVQOHDISGNKWL VYGA 60  
QY 61 PLETNQYQKTGDVYKCPVIHGNCTKLNLRVTLNSVSRKDNMELGLSLATNPKDNSFLA 120  
DB 61 PLETNQYQKTGDVYKCPVIHGNCTKLNLRVTLNSVSRKDNMELGLSLATNPKDNSFLA 120  
QY 121 CSPWASHECCSSYTTTCMCVRVNSNFRFSKTAPALORCQTTMDIVIVLQSGNSIYPWVE 180  
DB 121 CSPWASHECCSSYTTTCMCVRVNSNFRFSKTAPALORCQTTMDIVIVLQSGNSIYPWVE 180  
QY 181 VQFLINILKCFYIGPGQIQGVVVOYGEDVVVHPEHLNDYRSVKDWEAAASHIEQGGTET 240



```
Db 181 VQHFLLNLLKYYIGPQIQVGVVQYGEDVWHEFLNDYRSVKDVEAASHIEQSGTET 240
Qy 241 RTAFGIEFARSAFOKGRGKAKKVMIVITDGHSDSPLEKVIQOOSERDNVTRVAVL 300
Db 241 RTAFGIEFARSAFOKGRGKAKKVMIVITDGHSDSPLEKVIQOOSERDNVTRVAVL 300
Qy 301 GYNNRRGINPFTFLNEIKYIASDPDDKHFFNVTDAAALKDI VDALGDRIFSLEGTNKNET 360
Db 301 GYNNRRGINPFTFLNEIKYIASDPDDKHFFNVTDAAALKDI VDALGDRIFSLEGTNKNET 360
Qy 361 SFLGMSQTFSSHVVEDGVILGAVGAYDWNAGVILKETSAGKVIPLRESYLKBPPEELK 419
Db 361 SFLGMSQTFSSHVVEDGVILGAVGAYDWNAGVILKETSAGKVIPLRESYLKBPPEELK 419

RESULT 15
US-09-918-715-250
; Sequence 250, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Brad St. Croix
; APPLICANT: Bert Vogelstein
; APPLICANT: Kenneth Kinzler
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00134
; CURRENT APPLICATION NUMBER: US/09/918,715
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/222,599
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/224,360
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 250
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-918-715-250

Query Match 34.8%; Score 2166.5; DB 12; Length 1179;
Best Local Similarity 38.1%; Pred. No. 4.9e-190;
Matches 464; Conservative 239; Mismatches 425; Indels 89; Gaps 22;

Qy 4 PR---GLVVA--WALLSLWPGFTDTFNMDTRKPRVPGSRTAPPGYTVVOHDLISGNKVLV 58
Db 5 PRARPGVAVACMLLTVLVLRCCVSPNVDVKNSTFSGPVEDMFGYTVQOYNEBQWVLI 64
Qy 59 GAPLETNGYQKTDGVYKCPVIHGN---CTKMLG-RVTLNSVSEKDNKRLGSLATNPK 114
Db 65 GSPLVGOPKRTGVDYKCPVGRGESLPCVKLDLPVNTSIPNVTVEKNKTFGSTLVNTP- 123
Qy 115 DNSFLACSLPMHEGSSYTTGMCSSRVNSNTRFSKTVPALQRCQTMDIVILVDGNS 174
Db 124 NGGFLACGLYAYRCGLHYTTGICSDVSPTQVNVSNAP-VQECSTQLDIVILVDGNS 182
Qy 175 IYVWVEVQHFLLNLLKXYIGPQIQVGVVQYGEDVWHEFLNDYRSVKDVEAASHIEQ 234
Db 183 IYVWVSVAFLNDLXRMDIGPKQTVQVGVQGENVWHEFLNLYSSSTEELVAAKKIVQ 242
Qy 235 RGGTTRTAFAGTIEFARSAF--OKGGRGAKKVMIVITDGRSHDSPLEKVIQOOSERDNV 292
Db 243 RGGRTMTALGTDTARKEAFTARGARRGVKKVMIVITDGRSHDNRLKVKVQDCEDENI 302
Qy 293 TVAVAVILGVYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDI VDALGDRIFSL 352
Db 303 QPFSFALLGSGYNNRNLSTKFEVEEIKSTASEPETEKHPFNVSDELALVTIVTKLGERIFAL 362
Qy 353 EGT--NKNETSFGLEMSQTGSSSHVVEDGVLLGAVGAYDWNAGVILKETSAGKVIPLRESYL 411
Db 363 EATADQSAASFEMEXSQTFSAHYSQDWMVLGAVGAYDWNAGVILKETSAGKVIPLRESYL 422
```

```
Qy 412 KEFPEELKNGHAYLGYTWTTSVVVSQRQGRVYVAGAPRPNHIGKVLFTMHNRSILTIHQAM 471
Db 423 VESTKNEPLASYLGYTVNGSATAGSDVLYIAGQPRNHTIQVLIYRMEDG-NIKILQTL 481
Qy 472 RGOIGSYFGSEITISVDIDGCVGTDVLLVGLAPMTF-NEGRERGKVVYVVELQNEFYNGT 530
Db 482 SGEQIGSYFGSILTTTIDKDSNTDILLVGLAPMTYMGTEKEBQGVYVVALNOTFPEYQMS 541
Qy 531 LK-----DSHSYQN-----ARFGSSIASYVRDLNQDSYNDVNVVVGAPLEDNEHAG 572
Db 542 LEPIKQTCSSRQNSCTTENKNEPCGARGTAIAAVKDLNLDGFNDIVIGAPLEDDHGG 601
Qy 573 AIYTFHGRGSIKTPQORITASBLATGLQVFGSGISGQLDNLDELGLDLAVGALGNVAVI 632
Db 602 AVYIYHSGSGKTIRKEYAQIRPSGGDGKTLKFPGSGIRGEMDLNGDGLTDTVTIGLGGAL 661
Qy 633 LMSRPVQVQINASHLFEPKINIHRDCKRSORDATCLAAFLCTPIPLAPHFQTTTVGIR 692
Db 662 FWRDVAVKVVMFEPKNIQKONCHMEKFTVCINATVCFEVKLKSKEDTIEADLQ 721
Qy 693 YNATMDERRYTPRAHLDGEGDRFTNRVAVLSSGGSELCEINHFVLTADYVYKPTFSVEY 752
Db 722 YRVTLDSLRLQISRSFFSGTOERKVQRNITVRKSE--CTGHSFYMLDKHDFQDSVRITLDF 779
Qy 753 SLEDPDHGPMDDGWPPTTLRVSVPPWNGCNEDEHCVPLDLVLDARSULTAMEYCORLKR 812
Db 780 NITDPENGPVLDDSLPNSVHEYIIPFAXDCGNKCKISDL----- 818
Qy 813 PAQDCSAYTLISFDTTTFIESTRQRAVEATLENRGENAYSTVLNISQSANLOFASL--I 870
Db 819 -----SLHVATTEKDLIVRSQNDKFNVSILTAKNDSAYNRTIVHYSPLVFSGIEAI 873
Qy 871 QKEDSDGSIKCVWEERLQKQVCNVSYPFPRAKAKVAFRLDSEFSKSIPLHLEISLAAG 930
Db 874 QKDSCEGN-----HNITCKGYPELRRGEMVTFKILQFNTSYLKENVTIYLSAT 923
Qy 931 SDSNERDSTKEDNVAPLRFHLKYEADVLFTRSSLSHVEVKLNS-----LERYDQIGPFF 986
Db 924 SDSEPPETLSDNVVNISIPVKEYEVLQFYSSASEYHISIAANETVPEVINSTEDIGNEI 983
Qy 987 SCITRIQNLGLPIHGMKMITPIATRSNGRLKLRDPLTDEANTSC-----NIMGNS 1040
Db 984 NIFYLRKSGSPFPMPELKLSIFPNMTSNGYPVL-YPTGLSSSENANCRPHIFEDPFSIN 1042
Qy 1041 TEYRPTVVEEDLREAPOLNHSNDVVSINCNIRLVPNQEINPHILGLMLRSLKALKYKS 1100
Db 1043 SKGKTTSTDLKKGTTLDCTCKPATITCMLTSSDISQVNSLI--LWKPTPIKSYFSS 1100
Qy 1101 MKIMVNAALQRFHSPFIFREEDPSRQTEFISKQE-DWQVPIHVIIVGSTLGGLLALL 1159
Db 1101 LNLITIRGEL-RSENASLVSSNQRELAQISKDGLPGRVPLWILLISAPAGLLMLL 1159
Qy 1160 VLARKLGFPSARRR 1176
Db 1160 ILALWKIGPFRPLKKX 1176
```

Search completed: June 24, 2004, 18:02:34  
Job time : 58 secs







Db 248 DLITFGAIQYARKYASAGGRSATSATKVVVVVTDGSHDGLKAVIDQCNHDLIRF 307  
Qy 236 AVAVLYNNRRGINPE-TFLNEIKYIASDPDDKHFNFVTDAAALKDIVDALGDRIFSLG 354  
Db 308 GIAVLGLNRLNADTKMLIKSEIKALASIPTEYFFNFVSDAALLKXAGTLGEQIFSI 367  
Qy 355 INKNETSGLEMSQTGFSHV--VEDGVLLGAVGAYDNGAVLKEISAGKVIPIRESYLK 412  
Db 368 TVXGDNFXMMSQVGSADYSQNDILMLGAVGAGVSGTIVQKTSHEGLI----- 419  
Qy 413 EFPBEL-----KNHGAYLGYTVTSVSSROGRVVVAGAPRFNHTGVILFTMHNRSLT 466  
Db 420 -FPQAEDIOIDRESSVILGVV-AAISTGSHFVAGAPRANTGVQIVLVSVNENGIT 477  
Qy 467 HQMARGOQISYSGSEITSDIDGVDVLLVGAAPYFNE-GRERKGVVYVBLRNRF 525  
Db 478 VIOHRGQOISYSGSVLCSVDVDKDTITDVLVGAAPYMSDLKKEGKVVLFTIKGIL 537  
Qy 526 VYNGTLKXSHSYQVAFSSIASYDNDQSDVNDVVGAPLEDNAGAIYIFHCFRGSIL 585  
Db 538 GHOPELEGEGIENTRFGSALAAUSDINDGDFNVIVGSPLENQNSGAVIYNGHGTIR 597  
Qy 596 KIPKQIRITASBLA--TGLOYPGCSIHGOLDINEDGLDLAVGALGNVILMSRPVQINA 643  
Db 598 TKYSQKILGSDGAFRSHLOYPGCSLDGYDGLNGDSITDVSIGAFGVVQWLSQSIADVAI 657  
Qy 644 SLRPEPSKINIFHRDCKESGRDACLAFICFTPIELAPHQFTTVTGIRYNATWD----E 699  
Db 658 EASTPEKITILVNNAQ-----IILKLCFSKAP-RPTQNNQVAVIYVNTILDAGFS 708  
Qy 700 RRYTPRAHLDGGRDFTNRVALLSGQELCERINPHVLDITADYVVKPV--TFSVEYSLRDP 758  
Db 709 SRVTSRGLFKENNERCLQKNVVAQAQSCPEHI-IYIQEPSDVVNSLDDLRLVDISLENG 767  
Qy 759 HGPMLDDGWPTLRSVSPFWGNCHEDEHCVDPDLV--DARSLPTAMEYQORVLRKPAQC 817  
Db 768 TSPALAYSETAKVFSIPFHDCGEDGKCSIDLVLQDVR--IPAAQE-----QP---- 814  
Qy 818 SAYTSLFDTTFTIESTQRVAVETLENRGENAYSTVLNISQSANLQFASLIQKEDSDG 877  
Db 815 -----FIVSNQNKRLTFSTVLNKRKESAYNTGIIVDFSENFFASFLPVD--- 860  
Qy 878 SIECVNEERIKQV--CNVSPFFRAKAVAFRLDSEFSKSIPLHLEIELAAGSDSNER 936  
Db 861 GTEVTQVAAQSKSVACDVGYPALKREQQVTFINDFNLQNLQQAASLSFQALSESQEE 920  
Qy 937 DSTKEDVAPLRFHLKYEADVLFTSSSLSHYEV----KLASSLERYDGIQPPFSCIFRI 992  
Db 921 N--RADNLVNLIKIFLYDAEHLFRTSTNINFTISSGNGVPSIVHSPEDVGPKE--IFSL 976  
Qy 993 Q-NLGLFPIHGMKMITIPIATRSNRLKLRLDPLTDRANTSCHNIGNSTEYRTPV--- 1048  
Db 977 KVTGSGVPSMATVIIHIPOYTKENPLMYLTGVQTDKAGDIC---CNADINPLKIGOT 1032  
Qy 1049 -----EEDLRAPQLNHSNDVVSINCNRLVNPQINPHILG-----NLWLSL 1093  
Db 1033 SSSVSFKSENFRHTKELNCRATASCNVTCWLK-----DYHMKGEYFVNVVTRINWTF 1085  
Qy 1094 KALKYKMKIMWNAALQRFSPFFREDEFSRIEPEISK-QEDMOWPIWIIIVGSLIG 1152  
Db 1086 ASSIFQT--VQLTAAAEINTNPEIYVIEDNTVPIPLMKPDEKAEVPTGVIGSLIAG 1143  
Qy 1153 LLLALLVALRKLGFRF 1170  
Db 1144 ILLALLVALRKLGFRF 1161

## RESULT 3

US-08-485-618-53  
; Sequence 53, Application US/08485618  
; Patent No. 5728533  
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Garstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,618  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: US 08/173,497  
APPLICATION NUMBER: 23-DEC-1993  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32797  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 53:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1161 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-485-618-53

Query Match 19.2%; Score 1194.5; DB 1; Length 1161;  
Best Local Similarity 30.1%; Pred. No. 2.3e-98;  
Matches 374; Conservative 198; Mismatches 507; Indels 165; Gaps 49;

Qy 5 RGLVW---AWALSLWPGFTDTFMMDTRKPRVPGSRTAPRGYTVQQHDISGNKWLAVGAP 61  
Db 3 RGVVILCGWALASCHG---SNLDVEKP-VVFKDAASFGQTVVQP---GGSLVVGAP 54  
Qy 62 LETNGYQKTGDVYKCPVIHGNCTKLNIGRVTLNVSRKDNMRGLGLSLATNPKNDSFLAC 121  
Db 55 LEAVAVNQTSQSDCCPACVCPILL-HIPLAV-----NMSLGLSLVADTNNSQLAC 108  
Qy 122 SPVWSHECGSYTYTGMCSRNVNSNFRPSKTVAPALORCQ-TYMDIVIVLQGSNSI--YFW 178  
Db 109 GPTAQACAKMVAKSGCLLGLSSLOFQAIPATMPECPCQEMDIAFLIDGSGSIDQSF 168  
Qy 179 VEVQHFLNLIKXFIYIGPGQI-----QVGVOVGEDVHFHFLNDYRSVXDVVEAASHIE 233  
Db 169 TQMKDFVYKALM-----GQLASTSTSFSLMQYSLNLIKHTFTTEPKSSLSQSLVDAIV 221  
Qy 234 QRGQTEFTRTAFGIEFARSEAP--QKGRKGAKKVMIVITDGEHDSF-DLEKVIQOBERD 290  
Db 222 QLQGL-TYTAGIQKVVKELFHSKNGARKAKILIVITDQKPRDPLEYRHPVPEAKA 280  
Qy 291 NVTRYAVAVLGYNNRGINPETFLNELKYIASDPDDKHFNFVTDAAALKDIVDALGDRIF 350  
Db 281 GIIRYAIQVGDAPRE-----PTALQELNTIGSAPSDHVFVKGVFVALRSIQRIQIEKIF 335

```

351 SLEGT-NKNETSPGLEMSOTGFSHVVDGVLGAVGAYDNGVNLKETSAGKVIPLRES 409
336 AIEGTSSSSPHEMSQGFSALEMDGVPVLCVAVGFSNGCAFIPY-----NMRST 390
410 YLKEPPEELKHGAYLGYTVTSVSSQGRVYVAGAPRPNHTGKVLFTMNNRSLTHQ 469
391 PINMSQENEMDRDAYLGS-TALAFWKGVHSLILGAPRHQHTCKVIFT-QESRHWPKS 448
470 AMRGOQIGSYFGSEHTSDVDGQVTDVLVAGAPMFNEGRCKVYVYELR--QNRVY 527
449 EVRTQIGSYFGSGLSCVSDMDRGSTDLVLIGVPHYEHTR--GGQVSCVCPMPGVRSMHC 507
528 NGTLKDSHYNARFPGFSIASVRDLNQDSYNDVVVVGAPLEDNAGAIYIFHG-FRGSILK 586
508 GTTLRGEGQHPWCFGAALTVDGVDNGDSLADVAIGAPGEENKAVIIFHGASRQDIAP 567
587 TPQRITASLALGLOVFGSIIHQDLNEDGLIDLAVGALGNVILMSPVVOINASLH 646
568 SPQRVTGSLFLRLQVFGSLGGQDLTQGLDVLAVGAQGHVLLRSLLPLKVGISIR 627
647 FEPKINIFHRDCKSRDATCL--AAFLCFT-----PIFLAPHFTTTVGVRYNATMD 698
628 FAFSEVAKTVQC--WGRTPVLEAGEATVCLTVKGSPLDG---DVQSSVRYDLALD 681
699 ERYTPRAHLDGGRDTRNRAVLLSQELCERINFRVLD-TADYKVPVTFVSVEYSL-- 754
682 PGRIGRAIPDETCKNCTLRKTKLGLGDH-CETMKULLPDCVEDAVTPILRLNLSLAGD 740
755 EDPDHG--PMLDDGWPPTLVSVVPFWNGCNEDEHCVPLDVLDAESDLPTAMEYCORVLRK 812
741 SAPSRNLFPVLAVSQDHVASTVFPFKNCKQELLCEGNL-----779
813 PAQCSAYLSPFTTVPILSTQRVAVEATLENGENAYSTVINISQSANLOPASLIQK 872
780 -----GVSNFSGLOVLEVSSPBLTVTVVWNEGEDSYGTLIKFPYPAELSRYRVTRA 833
873 ESDG--STECVNE---ERLQKQVNVSPYFRAKAKVAFPLDSEFSKSIPL-HHLRI 925
834 QQEPYPLRLACAEPTGQSLRASSCSINHPFRGAKATMITFDVSKAPLGRLL 893
926 ELAAGSNERDSTKDNVAPLPHLYKADVLFTRSS-----SLSHYEVKLNSSLE 977
894 RASASSENKPEPSK-TAPQELPVKYTVTVISQEDSTKHFNFSSHGB-RQKEAHH 950
978 RYDGIGPFPCIFRIONLGLPPIHGMKMTIPIATRSNRL--LKLRLDPLDEANTSCN 1035
951 RY-----RVNLSPLT-AISVNFWPILL-NGVAVWMDVTLR---SPAQGVSC- 993
1036 IWGNSTERYRTPVEEDLRAPQ---LNHSNSDVWSINCM-RLVPMQEIHFHLLGNL-- 1088
994 ----VSQREPPQHSLLTQIQGRSVLDCAIADCLHLRCDIPSLGLDELDFILKGNLSF 1048
1089 -WLSL---KALKYKSKIMYNALQROPHSPITREBPSRQIBFEISKQEDWQVPIWI 1144
1049 GWISQTLQKVLILLSABITFNTSVYSQPLGQBAFLRAQVSTWLEBYVYVYB-----PVFL 1103
1145 IGVSTGLGMLLALLVLRKLKGFRRARRRR---PGLDPTP 1184
1104 MVFSSVGGULLLALLIYALYKLGFFK--RQYKEMLDLPSPADPP 1145

```

```

RESULT 4
US-08-362-652-53
; Sequence 53, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower

```

```

; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1161 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-652-53

```

```

Query Match 19.2%; Score 1194.5; DB 1; Length 1161;
Best Local Similarity 30.1%; Pred. No. 2,3e-98;
Matches 374; Conservative 198; Mismatches 507; Indels 165; Gaps 49;

QY 5 RGLVV---AWALSMPGFTDFNMDTRXPRVIGSRTAFPGYTVQOHDISGNKVLVVGAP 61
DB 3 RGVVILCCMALASCHG---SNLDVEK2-VVFKEAASFGQTVVQF---GGSRLVVGAP 54
QY 62 LETNGYQKTVGVYKCPVTHGCTKLNLRGVLTVSNVSKDNKRLGLSLATFNPKDNLFLAC 121
DB 55 LEAVAVNGTQSSDCPPATGVCOPILL-HIPLAV-----NMSLGLSLVADTNNSQLLAC 108
QY 122 SPLMSHECGSSYTTTGMCSRVNSNFRFSKTVAPALQRCQ-TYMDIVIVLDSNSI--YFW 178
DB 109 GPTAQACAKNMYAKGSCLLGSSLOFTQAIPTATWPCPGQEMDIAFLIDGSGSIDQSD 168
QY 179 VEVQHFLNLIKFFVIGPGQI-----QGVVQYGEDVYVHEFLNDYRSVKDVVEAASHIE 233
DB 169 TOMKDFVKALM-----GQLASTSTFSLMQYINILKTHPTTFEKSLSLSPQSLVDALV 221
QY 234 QRGFTETRTAFGISEAFSEAF--QKGGKKGAKVMIVITDGHSDSP--DLEKVIQOSPRD 290
DB 222 QLOGL-TVTASGIQKVKELFHSKNGAKSAKLILVITDQKPRDPLEYRHRVPEAEKA 280
QY 291 NTRYAVAVLGYNRRGINPETFLNEKYIASDDDDKHFFNVYTDAAALKDVLADGDSIF 350
DB 281 GIIRYAVIGVDAFRE-----FTALOELNTIGSAPQDHVPKGVNFVALRSIQKQEKIF 335
QY 351 SLEGT-NKNETSPGLEMSOTGFSHVVDGVLGAVGAYDNGVNLKETSAGKVIPLRES 409
DB 336 AIEGTSSSSPHEMSQGFSALEMDGVPVLCVAVGFSNGCAFIPY-----NMRST 390
QY 410 YLKEPPEELKHGAYLGYTVTSVSSQGRVYVAGAPRPNHTGKVLFTMNNRSLTHQ 469
DB 391 PINMSQENEMDRDAYLGS-TALAFWKGVHSLILGAPRHQHTCKVIFT-QESRHWPKS 448
QY 470 AMRGOQIGSYFGSEHTSDVDGQVTDVLVAGAPMFNEGRCKVYVYELR--QNRVY 527

```

Db 449 EVRGTOIGSYFGASLCSDMDRGGSTDLVLIGPHYEHTR-GGQSVSCMPGVRSRWHC 507  
Qy 528 NGTLKDSHSHYQARPOSSIASVRLDNDSDYNDVWVVGAPLEDNEAGAVIFHG-FRGSILK 586  
Db 508 GTTLGEGQHPGRGAAITVLGVNDGSLADVAIGAPGEENRGAIVIFHGASQDIAP 567  
Qy 587 TPQRITASELATGQYFQCSHIGLDLNEGLDLAVGALGNVILMSRPVQINASLH 646  
Db 568 SPSQRTVSQLEFLQYFQCSLGGQDLTQDGLVLDVAVGAQGHVLLSLPLKVGISIR 627  
Qy 647 PEPSKINIPHRCKSGRDATA---AAELCT---PIFLAPHFQTTVGIIRNATWD 698  
Db 628 PAPSEVAKTYQC-WGRPTVLEAGEATVUTVVKGSDDLIG---DVQSSVRYDLALD 681  
Qy 699 ERYTPRAHLDEGGDRFTNRVALLSSGQELCERINFHVL-DADYVVKPTFSEYSL---754  
Db 682 PGLISRAIFDETKNCTLRRTKTLGLGDH-CETMKLLLPDCVEDAVTPILRLNLSLAD 740  
Qy 755 EDPDHG-PMLDDGHTTLRVSPVFWNGCNEDEHCVPLDVLDRSDLPAMEYCORVLR 812  
Db 741 SAPSRNLRLVAVGSDHVTAGFPPEKNCKQELLCEGNL-----779  
Qy 813 PAQDCSAYTSLPDTTTFIESTQRQVAVBATLENGENAYSTVLNISQSANLQFASLIQ 872  
Db 780 -----GVSNFSGQLVGVSGSPELTVTVVWNEGSDSYGLIKFYPAELSRYRTA 833  
Qy 873 EDSGD---SIECNE---ERRLOKQVNVSPFPRAKAVAFRLDSRFSKSIPL-HHLEI 925  
Db 834 QQPHYPLRLACEABPTGESLRSSCSINHPIFREGAKATEMITFDVSYKAFGLDRLL 893  
Qy 926 ELAAGSDSNERSTKEDNVAPLPHLYEADVLFRSS-----SLSHVEVLKNSLE 977  
Db 894 RASASENKPISK-TAFQLELPKVTYTVVISROEDSTKHFPSSHGE-ROKRAEH 950  
Qy 978 RYDGIPGPSCIFRIGNLGLFPIHGMMKXITIPATRSNRL-LKLRDFLTDEANTSCN 1035  
Db 951 RY-----RVNLSPLTL-AISVNFVPIILL-NGVAVMDVTLR---SPAQGVSC- 993  
Qy 1036 INGNSTERYPTVEEDLRAPQ-----LNHSNDVVSINCN-RVFNQBIHFLLGNL---1088  
Db 994 -----VSOREPPOHSDLLTQIGRSVLDCAIADCLHLRCDIPSLGTLDDELDFILKGNLSF 1048  
Qy 1089 -WLRLS---KALKYKSMKIMVNAALQROFSPFIFREEDPSRQIBEFISKQEDMOVPTWI 1144  
Db 1049 GWISQTLQKVLLESEAEITNTSVYSQLPGQEAFLRAQVSTMLEEYVYB-----PVFL 1103  
Qy 1145 IVGSLTGLLLILALVLRKLGFPSARRRE-----PCLDPTP 1184  
Db 1104 MVFSSVGGLLLALITVALYKLGPFK--ROYKEMLDLPSADPD? 1145

## RESULT 5

US-08-605-672-53  
; Sequence 53, Application US/08605672  
; Patent No. 5817515  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/605,672  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jz., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 53:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1161 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-605-672-53

Query Match 19.2%; Score 1194.5; DB 2; Length 1161;  
Best Local Similarity 30.1%; Pred. No. 2.3e-98;  
Matches 374; Conservative 198; Mismatches 507; Indels 155; Gaps 49;  
Qy 5 RGLVV---AWALSIMPQGTDTENNDRKPRVIPSRTAFPGYTVQOCHDISGNKMLVWGP 61  
Db 3 RGVVILLCGWALASCHG---SNLDVEKP-VVPKEDAAAPGQTVVQF---CGSRLVWGP 54  
Qy 62 LETNGYOKTGDVYKCPVTHGNCTKLNAGRTVLSNVSRKONMRLGLSLATWPKDNSFLAC 121  
Db 55 LEAVAVQGTQSSDCCPATGVQCPIIL-HIPLEAV-----NMSLGLSLVADTNNSQLAC 108  
Qy 122 SPLMSHBCSGSYTTGMCVRVNSNFRSKTVAPALQRCQ-TYMDIVLVLDGNSI--YWP 178  
Db 109 GPTAQACAKMVAKSGCLLGLSGLQIQAIPATMPCPGQEMDIAFLDGGSGIDQSDP 168  
Qy 179 VEVQHPFLNLIKXYIYGPQI-----QGVVVOYGEDVWHFHLNDYRSKDVVZAAASHIE 233  
Db 169 TOMKDFYKALM-----GQLASTSTFSLMQVSNILKTHFTTFEFKSSISPSQSLVDAIV 221  
Qy 234 QRGCTETRTAFGIEFARSEAF--QKGRKGAKKVMIVITDGESHDSF-DLEKVTQQSERD 290  
Db 222 QLQGL-TYTASGIQKVVKELFHSKNGARKSAKKLIVITDGKFRDPLEYRHHVPEAKA 280  
Qy 291 NVTRYAVAVLGYNNRGINPETFLNEIKYTASDDPKHPFNVTDEALKDIDVLDGRIP 350  
Db 281 GIIRYAIYGVGDAPRE-----PTALQELNTIGSAPSDHVFVKGVNFVALRSIQRIQEKIF 335  
Qy 351 SLEGT-NKNETSFGLMSQTFSSHVVEDGVLLGAVCAVDWNGAVLAKETSAKVIPLRES 409  
Db 336 AIEGTERSSSSSQHEMSQEGFSSALSNDGPVLGAVGFGFSGGAPLYPS-----NMRST 390  
Qy 410 YLKEFPPEELKNHGAYLQYTVTVSVSSRQGVVYVAGAFRNHTGKVLFTMNNRSLTIHQ 469  
Db 391 FINMSQENEDMRDAYLGYS-TALAFKGVHSLILGAPRHQHTGKVVIPT-QESRHWRFKS 448  
Qy 470 AMRQOQIGSYFGSEITSDIDGGVTDVLLVGAPMYENEGREGKVVVYELR--ONRFVY 527  
Db 449 EVRGTOIGSYFGASLCSDMDRGGSTDLVLIGPHYEHTR-GGQSVSCMPGVRSRWHC 507  
Qy 528 NGTLKDSHSHYQARPOSSIASVRLDNDSDYNDVWVVGAPLEDNEAGAVIFHG-FRGSILK 586  
Db 508 GTTLGEGQHPGRGAAITVLGVNDGSLADVAIGAPGEENRGAIVIFHGASQDIAP 567  
Qy 587 TPQRITASELATGQYFQCSHIGLDLNEGLDLAVGALGNVILMSRPVQINASLH 646



```

QY 699 ERYTPRAHLDGSDRFTNRAVLSSQELCERINPHVLD-TADYVVKPTFVSVEVL--- 754
Db 682 PGLISRAIFDFTKCTLTRKTLGLGDH-CETMKLLPDCVEDAVTPIILNLNLSLGD 740
QY 755 EDPDHG--PMLDDGWPPTILRVSPFWMGCEDEHCVPLVLDARSDLPAMEYQORVLRK 812
Db 741 SAPSRNLRPVAVGSDHVTASFPPEKNCKQELLCEGNI----- 779
QY 813 PAQDCSAYTSPDTTFFIESTRQAVAEATLENGENAYSTVLNISQSANLQFASLIQK 872
Db 780 -----GVSNFNSGLQVLEVGSPPELTVTVTVMNEDGDSYGLTIKFFYPABELSVRVTRA 833
QY 873 EDSG--SIBCVNE---ERRKQKQCVNVSPPFRKAKVAFRLDSEPSKSIPL-HHLEI 925
Db 834 QQPHYPLRLACEAPTGOESLRSSCSINHPIFREGAKATFMTEDVSYKAFGLDRLL 893
QY 926 ELAAGSDNERSTKEDNVAPLRLKYEADVLFTRSS-----SLSHYEVKLNSSLE 977
Db 894 RASASSENKPEISK--TAFQLELPVKYTVTVVISRQEDSTKHFNFSSSHGE-RKQEAH 950
QY 978 RYDGIGPPFCIFRIONGLPFIHGMKMTIPIATRSNRL--LKLDFLTDEANTSCN 1035
Db 951 RY-----RVNLSPLTL-AISVNFVPIILL-NGVAVMDVTLR---SPAGGVSC- 993
QY 1036 IWNSTYRPTVPEEDLRAPQ-----LNHNSDVVSINCN-RELVPNOEINFHLLGNL-- 1088
Db 994 -----VSQREPPQSHDILTOIGRSVLDCAIDCLHLRCDIPSLGTLDELDFILKGNLSP 1048
QY 1089 -WLRLS---KALKYKSMKIMVNAALORQHPFPFREDPSRQIEFELISKQEDQVWI 1144
Db 1049 GWISQTLQKVLLESEBITNTISVYSQLEPQAEFLRAQVSTMLEEYVYVYB-----PVFL 1103
QY 1145 IVGSTGLGLLLALLVLAIRKLGFRRSARRRE-----PGLDPTP 1184
Db 1104 MVFSSVGLLLALLITVALYKLGFK--ROYKEMLDLPSADPDP 1145

```

RESULT 7

```

US-08-943-363-53
; Sequence 53, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,363
; FILING DATE:
; CLASSIFICATION: 530
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.

```

```

; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1161 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-943-363-53

```

```

Query Match 19.2%; Score 1194.5; DB 2; Length 1161;
Best Local Similarity 30.1%; Pred. No. 2.3e-98;
Matches 374; Conservative 198; Mismatches 507; Indels 165; Gaps 49;

```

```

QY 5 RGLVV---ANALSLWPGFTDTENMDTKPRVIFGSRTPAGYTVQOHDISGNKWLWVGAP 61
Db 3 RGVVILLCGHALASCHG-----SNLDVEK-VFKEDAAAPGQTVVQF---GGSSLVVGAP 54
QY 62 LETNGYQKTDGVYKCPVIHGNCTKLNLRVTLNVSERKDNMRGLSLATNPKNDSFLAC 121
Db 55 LEAVAVNQTCQSSDCPEATGVCQPIIL-HIPLEAV-----NMSLGLSLVADTNNSQLAC 108
QY 122 SPLWSHECGSSYTTGMCVRNENFRPSKTVAPALQRCQ-TYNDIVIVLGGNSI--YFW 178
Db 109 GPTAQARACAKMYAKSCLLGLSSLOFIQAIPTAMPSCPGQEMDIAFLIDGSGSIDQSD 168
QY 179 VEYQHFLINILKFFYICPQCI-----QVGVVQYQGVQEDVHVEFHLNDRSVKDVVEAASHIE 233
Db 169 TOMKDFVKALM-----CQLASTSTSFSLMQSYNLKTHFTTEPKSSLSQSLVDV 221
QY 234 QRGCTETRTAFGIEFARSEAP--OKGRGAKKMTIVITDGEHSDSP-DLSKVIQOSERD 290
Db 222 QLOGL-TYTAGSIQRYVKELFHSKNGARKSAKILIVITDQGRFDPLEHYHVIPEAKA 280
QY 291 NVTRYAVAVLGYNNRRGINPETFLNRIKYIASDDPKHPFNVTDEAALXDIDVALGRIF 350
Db 281 GIIRYALGVGDARE-----PTALQELNITGSPSQDHVKGNFVALRSIQRIQKIF 335
QY 351 SLEGT-NKNETSPGLEMSQTFSSHVVBDVLGAVGVYDNGAVLAKETSAGKYIPIRES 409
Db 336 AIBGTSSRSSSSFQHEMSQEGFSSALSMDBPVLGAVGVGSWGGATLPS-----NMRST 390
QY 410 YLAKFPEELKNHGAYLGYTVTSVSSRQGRVYVAGAPRPNHGTGKVLFTMHNNRSLTIHQ 469
Db 391 FINMSQENEDMDRAYLGIS-TALAFWKGVHSLILGAPRHQHTKVVIFT-QESRHRPKS 448
QY 470 AMRGOQIGSYFGSEITTSVDIDGDVTDVLLVGVAPMYFNEGRERGRKYVYVELR--QNRVY 527
Db 449 EVRGTQIGSYFGASLCSVDMDRDGSTDVLVIGVPHYEHTR--GGQSVSCVPMGVRSRWHC 507
QY 528 NGTLKDSHYSQXARFGSSIASVEDLNODSYNDVWVQCAPLEDNHAGAIYIPIHG-FRGSILK 586
Db 508 GTTLHGEQGHWPGRFGAALTVLGDVNGDSLADVAIGAPGEEENRGAVYIPIHGASRDIA 567
QY 587 TPKQRITASELATGLQYFGCSIHGOLDNEDGLIDLAVGALGNVILWSPVVOINASLIH 646
Db 568 SPQRVTGSQLFLRLQYFGCSLSGGQDLTQDGLVDLAVGAQGHVLLRLSLPLKVGISIR 627
QY 647 FEPKINIHRDCKSGRQATCL---AAPLCFT-----PIFLAPHEQTTTGVIRYNATMD 698
Db 628 FAPSEVAKTVYQC--NGRTPTVLEAGATVCLTVARKGSPDLLG-----DVQSSVRYDLAD 681
QY 699 ERYTPRAHLDGSDRFTNRAVLSSQELCERINPHVLD-TADYVVKPTFVSVEVL--- 754
Db 682 PGLISRAIFDFTKCTLTRKTLGLGDH-CETMKLLPDCVEDAVTPIILNLNLSLGD 740
QY 755 EDPDHG--PMLDDGWPPTILRVSPFWMGCEDEHCVPLVLDARSDLPAMEYQORVLRK 812
Db 741 SAPSRNLRPVAVGSDHVTASFPPEKNCKQELLCEGNI----- 779

```



QY 813 PAQCSAYTSLPDTTTFIIBSTRQVAVAEATLENROENAYSTVLNISQANLOFASLIQK 872  
 Db 780 -----GVSNFSGLOVLEVGSSPELTVTVTWNEGEDSYGLIKFYPAELSYRVRTRA 833  
 QY 873 EDSG---STECVNE---ERRLOKQVCNVSYPFRAKAKVAFRLDSEPSKSIPL-HHLEI 925  
 Db 834 QQPHYPLRLACEAEPTGOESLESSCSINHPIFREGAKATMTITPDVSYKAFGLDRLLL 893  
 QY 926 ELAAGSDSNERSTKEDNVAPLRFHLKYEADVLFTRSS-----SLSHYEVKLSNLSLE 977  
 Db 894 RASASSENKPKTSK--TAPQLELPVKYTVTVVISRQEDSTKHFNFSSSHGE--RQKEAEH 950  
 QY 978 RYDGIGPPPSCIFRIQNLGLFPIHGMWKITIPATRSNRL--LKLDRFLDTEANTSCN 1035  
 Db 951 RY-----RVNLSPLTL-AISVNFVWPILL-NGVAVDVTILR-----SPAQGVSC- 993  
 QY 1036 IWGNSTEYRPTVEEDLRAPO-----LNHSNDVVSINCNI-RLVNPQEIHFHLLGNL-- 1088  
 Db 994 -----VSQREPPQHSLLTQIQGRSVLDCAIDCLHLRCDIPSLGTLDELDFILKGNLSF 1048  
 QY 1089 -WLRLS---KALKYKSMKIMVNAALQRPSPFIREDPSPQIEFESKQEDWQVPIWI 1144  
 Db 1049 GWISQTLQKVKLLLSABEITFTSVYSQLPQGEAFILRAQVSTMLSEYVYVE-----PVFL 1103  
 QY 1145 IVGSTLGGILLALLVLALRKLGPFRSARRRE-----PGLDPTP 1184  
 Db 1104 MVFSSVGGILLALLITVALYKLGFFK--ROYKEMLDLPSPADPDP 1145

## RESULT 8

US-09-193-043-53

; Sequence 53, Application US/09193043

; Patent No. 6251395

; GENERAL INFORMATION:

; APPLICANT: Gallatin, Michael W.

; TITLE OF INVENTION: No. 6251395el Human 2

; FILE REFERENCE: 27866/35004

; CURRENT APPLICATION NUMBER: US/09/193,043

; CURRENT FILING DATE: 1998-11-16

; EARLIER APPLICATION NUMBER: 08/173,497

; EARLIER FILING DATE: 1993-12-23

; EARLIER APPLICATION NUMBER: 08/286,889

; EARLIER FILING DATE: 1994-08-05

; EARLIER APPLICATION NUMBER: 08/362,852

; EARLIER FILING DATE: 1994-12-21

; EARLIER APPLICATION NUMBER: 08/943,363

; EARLIER FILING DATE: 1997-10-03

; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 53

; LENGTH: 1161

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-193-043-53

## Query Match

Best Local Similarity 30.1%; Score 1194.5; DB 3; Length 1161;

Matches 374; Conservative 198; Mismatches 507; Indels 165; Gaps 49;

QY 5 RGLVV---AWALSUWFGTDFNMDTRKPRVPGSRTAFFGYTVQOHDISGNKWLWVGAP 61  
 Db 3 RGWILLCGWALASCHG---SNLDVEKP-VVFKEAASFGQTVVQF---CGSRLWVGAP 54  
 QY 62 LETNGYQKTVGYKCPVTHGKCTKLNLRVTLSNVSERKNNRGLSLATNPKDNLPLAC 121  
 Db 55 LEAVAVNOTGSSDCPPATGVCQPIILL-HIPEAV-----NMSLGLSLVADTNNNSQLLAC 108  
 QY 122 SPLMSHECGSSYYITGMSRVSNNFRFSKTVAPALQRCQ--TYMDIVIVLQGSNSI--YPW 178  
 Db 109 OPTAQRACAKNMYAKSGCLLIGSSLOFTQOIPATMPECPQGEQMDIAFLIDSGSIDQSDF 168

QY 179 VEQHFELINILKXYIIGPQOI-----QGVVVQYGEDVVHFLNDYRSKVDVVEAASHIE 233  
 Db 169 TOMKDFVKALM-----QOLASTSTSFSLMOYSNILKTHFTTFEPKSSLSQSLVDAIV 221  
 QY 234 QRGTEHRTAFGIEPASEAP--QKGGKKGAKKVMIVITDGHSDSP--DLEKVIQOOSRD 290  
 Db 222 QLOGU-TYTAGGIQKVVKELFHSKNGARKSAKKILIVITDQCKPRDPLRYHVIPEAKA 280  
 QY 291 NVTRYAVAVLGYVNRGINPETFLNEIKYIASDPDDKGFNFVNTDEAALKDIDALGDRIF 350  
 Db 281 GIIRYALGVGDAPRE-----PTALQELNTIGSAPSQDHVFVKGVFNALLRSIQRIQEKIF 335  
 QY 351 SLEGT-NKUNTSFGLMSQCFSSSHVVEDGVLLGAVGAYDNGAVLXKTSAGKVIPIRES 409  
 Db 336 AIBGTESRSGSSFOHMSQGFSSALSNDGPVLGAVGFGSSGAGFYVPS-----NMKST 390  
 QY 410 YLKEFPBELKNHGYLAYLYTIVTVSVSSRQGRVYVAGAPRFNHTGRTKRVILFTMHNNRSLTIHQ 469  
 Db 391 FINMSQENEDMRDAYLGYS--TALAFWKGVHSLILGAPRHQHTGKVWIFT--QESSHWPKS 448  
 QY 470 AMRCQOIGSYGSEIITSDVDIGDQVTVLLVAGAPMYNEGRERKQVTVYELR--QNRPVY 527  
 Db 449 EVRGTQIGSYFGASLCSDMDRDGSTDVLVIGVPHYHYEHTR--GGQVSVCPMPGVRSWHC 507  
 QY 528 NGTLKDSHSYQNAFFGSSIASVRDLNODSYNDVVVVGAPLEDNHAGAIYIFHG--FRGSILK 586  
 Db 508 GTLLHGEQGHWPGRFGAALTVLGDVNGDSLADVAIGAPGEENREGAVVIFHGASRQDIAP 567  
 QY 587 TPQRITASBLATGLQYFGCSIHQDLNEDGLIDLAGALGNVILWSRPPVQVAINSLH 646  
 Db 568 SPQRVTGSQLFLRLQYFGQSLSGGQDLTQDGLVDLAAGQGHVILLRSLPLKVGISIR 627  
 QY 647 FEPSKINIFHRDCKRSGRDATCL---AALFLCT---PIFLAPHFQTTVTGIRYNATMD 698  
 Db 628 PAFSEVAKTVQC--WGRTPVLEAGEATVCLTVRKGSPDLG---DVQSSVRYDLALD 681  
 QY 699 ERRYTPRAHLDGGDRFTNRAVLLSSQOELCEINHFVLD--TADYVVPVTFSEYSL---754  
 Db 682 PGLRISRAIFDETNNCTLTRRKTGLGDH--CETWKLPLPCDVEDAVTPIILRLNLSLAD 740  
 QY 755 EDPDHG--PMLDDGWPTTLRVSVFPWNGCNEDEHCVPLDVLDAESDLPAMEYQORVLRK 812  
 Db 741 SPSRNLRPVLAVGSQDHVTASTPFPEKNCKQELLCEGNL-----779  
 QY 813 PAQCSAYTSLPDTTTFIIBSTRQVAVAEATLENROENAYSTVLNISQANLOFASLIQK 872  
 Db 780 -----GVSNFSGLOVLEVGSSPELTVTVTWNEGEDSYGLIKFYPAELSYRVRTRA 833  
 QY 873 EDSG---STECVNE---ERRLOKQVCNVSYPFRAKAKVAFRLDSEPSKSIPL-HHLEI 925  
 Db 834 QQPHYPLRLACEAEPTGOESLESSCSINHPIFREGAKATMTITPDVSYKAFGLDRLLL 893  
 QY 926 ELAAGSDSNERSTKEDNVAPLRFHLKYEADVLFTRSS-----SLSHYEVKLSNLSLE 977  
 Db 894 RASASSENKPKTSK--TAPQLELPVKYTVTVVISRQEDSTKHFNFSSSHGE--RQKEAEH 950  
 QY 978 RYDGIGPPPSCIFRIQNLGLFPIHGMWKITIPATRSNRL--LKLDRFLDTEANTSCN 1035  
 Db 951 RY-----RVNLSPLTL-AISVNFVWPILL-NGVAVDVTILR-----SPAQGVSC- 993  
 QY 1036 IWGNSTEYRPTVEEDLRAPO-----LNHSNDVVSINCNI-RLVNPQEIHFHLLGNL-- 1088  
 Db 994 -----VSQREPPQHSLLTQIQGRSVLDCAIDCLHLRCDIPSLGTLDELDFILKGNLSF 1048  
 QY 1089 -WLRLS---KALKYKSMKIMVNAALQRPSPFIREDPSPQIEFESKQEDWQVPIWI 1144  
 Db 1049 GWISQTLQKVKLLLSABEITFTSVYSQLPQGEAFILRAQVSTMLSEYVYVE-----PVFL 1103  
 QY 1145 IVGSTLGGILLALLVLALRKLGPFRSARRRE-----PGLDPTP 1184  
 Db 1104 MVFSSVGGILLALLITVALYKLGFFK--ROYKEMLDLPSPADPDP 1145

## RESULT 9

US-09-688-307A-53  
; Sequence 53, Application US/09688307A  
; Patent No. 6432404  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 6432404el Human Beta-2  
; FILE REFERENCE: 27866/36646  
; CURRENT APPLICATION NUMBER: US/09/688,307A  
; CURRENT FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 09/193,043  
; PRIOR FILING DATE: 1998-11-16  
; PRIOR APPLICATION NUMBER: 08/605,672  
; PRIOR FILING DATE: 1996-02-22  
; PRIOR APPLICATION NUMBER: 08/173,497  
; PRIOR FILING DATE: 1993-12-23  
; PRIOR APPLICATION NUMBER: 08/286,889  
; PRIOR FILING DATE: 1994-08-05  
; PRIOR APPLICATION NUMBER: 08/362,652  
; PRIOR FILING DATE: 1994-12-21  
; PRIOR APPLICATION NUMBER: 08/943,363  
; PRIOR FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 53  
; LENGTH: 1161  
; TYPE: PR1  
; ORGANISM: Mus musculus  
US-09-688-307A-53

Query Match 19.2%; Score 1194.5; DB 4; Length 1161;  
Best Local Similarity 30.1%; Pred. No. 2.3e-98;  
Matches 374; Conservative 198; Mismatches 507; Indels 165; Gaps 49;

QY 5 RGLVV---AWALSMPGTFDFTFMDTRKPRVPGSRATFGYTVQOHDISGNKMLVVGAP 61  
DB 3 RGVILLCCNALASCHG---SNLDVEKP-VVFEDAAAFQTVVQF---CGSLVLVGP 54  
QY 62 LETNGYQKTDVYKCPVHNGCTKLNIGRVTLNVSERKNRVLGLSLATNPKNDSFLAC 121  
DB 55 LEAVAVNQTQSSDPPATGVCQPIIL-HIPLAV-----NKSGLSLVADTNNSQLLAC 108  
QY 122 SPLSHGEGSSYYTTCMSRVNSFRSKTVAPLQRCQ-TYMDIVIVLDSNSI--YPW 178  
DB 109 GPTAQRACAKMYAKGSCLLGSSLQFTQAI-PATMPCPGQEMDIAFLDGSIGDQSF 168  
QY 179 VEQHFILNLIKFYIGPQOI-----QGVVQYGEDVVHFNLDNYSRVKDVVEAASHIE 233  
DB 169 TQKDFVKALM-----GQLASTSTSFSLMQYNSILKTHFTTPEKSSLSPQSLVDAIV 221  
QY 234 QRGTTETRTANGIEFAKSEAF--QKGRKGAKKVMIVITGESHDSP-DLEKVIQOGERD 290  
DB 222 QLOGL-TVTASGIQKVKELPHSKNGARKSAKILIVITDQKFRDPLEYRHVPEAKA 280  
QY 291 NVTYAVAVLGYNNRGINPETELNEIKYIASDPDDHFNVTDEAALKDIDVALGRIF 350  
DB 281 GIIYAVGVGDAFRE-----PTALQELNTIGSAPSDHVKGVNFVALRSIQOIKIP 335  
QY 351 SLEBT-NKNETSPGLEMSQTFSSHVEDGVLLGAVGAYDNMGAVLKETSGAKVPIRES 409  
DB 336 AIEGTESRSSSPQHEMSQEGFSSALMDGPFVLAGVGFSGGAFLYPS-----NMRST 390  
QY 410 YLKEPPELKNHGLVGYTVTVSVSRQGRVYVAGAPREHVTGKVLITMHNRSLIHQ 469  
DB 391 FINMSQENEDRDYLGYS-TALAFKGVHSLIUGAPRHQHTGKVIFT-QESGRHWRPKS 448  
QY 470 AMRQQIGSYRGSBITSDIDGDGVTDLVLLGAFMYFNEGRKGVVYELR--QNRVY 527  
DB 449 EVRGTOIGSYFGALSICVMDRDGSTDVLVIGVPHYEHTR-GGQVSVCPMPGVRSRWHC 507  
QY 528 NGTLKDSHSYQARFGSSIASVRLNODSYNDVWVGAPLEDNRAGAIYIIFHG-PRASILK 586

508 GTTLHCEGCHPWGRFGAALTIVLGVNDSLADVAIGAIPGBEENRGAVYIFHCASRQDIAP 567  
587 TPQORITASELATGLQYFGCSIHQDLNEDGLIDLAVGLGNNAVILWSRVVQINASLH 646  
568 SPSQRTGSQLFLRLQYFGCSLGGQDLTQDGLVDLAVGAGCHVLLRLSLFLKVGISIR 627  
647 FEPSKINIPHRDCKRGRDATCL--AAFLCFT-----PIFLAPHFQTTTGVIRYNATMD 698  
628 FAPSEVAKTVQC--WGRTPVLEAGEATVCLTVRKSPDLG-----DVQSSVRYDIALD 681  
699 ERRYTPRAHLDEGGDRFTNRAVLLSSQOELCERINFHVLQ-TADYVVKPVTFSVBSYL--- 754  
682 PGLISRAIFDETKNCTLTRKTLGLGDH-CETWKLILLPCVDEDAVTPILRLNLSLAGD 740  
755 EDPDHG--PMLDDGMPPTLVSVFVWNGNEDEHCVPLDVLDAARSOLDPTAMEYQORVLRK 812  
741 SPSRNLRPVLAVGSDHQHTASFPFKNCKOELLCEGNL----- 779  
813 PAQDCSAYTLTSDFTTTFIIBSTRQVAVEATLENRGENAYSTVINISQSANLOFASHIQK 872  
780 -----GVSNFSLQLVLEVGSSPELTVTVTVWNEGEDSYGTLIKFYVPAELSYRRVTRA 833  
873 EDSHG---SIECVNE---ERRLOKQVGNVYPPFRKAKVAFRLDSEFSKIFL-HHLEI 925  
834 QQPEPYPLRLACEAETPQESLSSSCSINHPFREGAKATFMTTFDVSYKAFGLGDRLLL 893  
926 ELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFTRSS-----SLSHVEVKLNSLE 977  
894 RASASSENKPEYSK-TAFQLELPVKYTVTVTVISQEDSTKHFNFSSHGB-RQKEAHE 950  
978 RYDGIOPPPFCIFRIQNLGLFPIHGMKMTIPIATRSNRL--LKURDPLTDEANTSCN 1035  
951 RY-----RVNLSPLTL-AISVNFVWPILL-NGVAVMDVTLT---SPAQGVSC- 993  
1036 IWGNSTYRTPVEEDLRRAPQ---LNHSNSDVWSINCNI-RLVPMQENFHLGHL-- 1088  
994 -----VSQREPPHSDLLTQIQERSVLDCAICHLRCDIPSLGTLDELDFILKGNLSP 1048  
1089 -WLRSL---KALKYKSMKIMVNAALQRFHSPIFREEDPSRQIEFEISKQEDQVPIWI 1144  
1049 GWISQTLQKXVLLSEAITFNTSVLSQPGQAFRLAQVSTMLEEYVVE-----PVFL 1103  
1145 IVCSTIGGLILLALLVLRKLGCFERSARERRE-----PGLDPTP 1184  
1104 MFSVSGVGLLLALITVALYKLGFFK--RQYKEMLDLPSADPDP 1145

RESULT 10  
US-09-350-259-53  
; Sequence 53, Application US/09350259  
; Patent No. 6620915  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 6620915el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350,259  
; CURRENT FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: 09/193,043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 53  
; LENGTH: 1161  
; TYPE: PR1



[illegible]

Db 576 SLSLRLQYFGQSLGGQDITQDGLVDLAVGAQHVLRLSLPLKVKVLSIRPAPMEVAK 635  
 Qy 655 FHRDCRSGRDATCL---AAFLCFTPIFLAPHFQTTTVG-IRYNATMDERRYTPRAHLDE 710  
 Db 636 AVQCWE--RTPTVLEAGEATVCLTVHKGSFDLLGNVQGSVRYDLALDPCRLISRIFDE 693  
 Qy 711 GDRFTNRVAVLLSSGQELCERINFHVL-DADYVVKVTFVSVEYL-----EDPDHGMPLD 764  
 Db 694 TKNCTLTGRKTLGLGDH-CETVKLLLPDCEVDAVSPILRLNLSLRVDSASPRNLHPVLA 752  
 Qy 765 DGNPTTLRVSVFPMWNGCNDEHCVDPDLVDARSDELPTAMEYCORVLRKPAQDCSAYTLSF 824  
 Db 753 VGSQDHIATSLPFKCKQELLCEGDL-----GISFNF 785  
 Qy 825 DTTVTIESTRORVAVATLENGENAYSTVLNISQSANLQFASLI---QKEDSDGSIEC 881  
 Db 786 SGLQVLVVGSPSLTIVTVWNEGEDSYGLTVFYPYAGLSYRVTGTQOPHYPLRLAC 845  
 Qy 882 VNE---ERRLQKQVNCVSYFFRAKAKAVAFRLDSEFSKIFL-HHLEIELAAGSDSNERD 937  
 Db 846 EAEPAQEDLRSSSCSINHPIFREGAKTTFMIFDVSKAFGLDRLLRKAKASSENKPD 905  
 Qy 938 STKEDNVAPLRFHLKYRADVLFTRSSLSHYEVLKNSLSRYDGIQPPFSCIPIRQNLGL 997  
 Db 906 TNK--TAFQLELPVKYTVYTLISRQEDSTNH-VNFSS---HGRROEAAHRYRVNLS 959  
 Qy 998 FEIHHMMKITIPIATRSQNLRLKLRDLFTDANTSCNIWGNSTERYPTVEE-----DL 1052  
 Db 960 LKL-AVRVNFVPLV-----LNGVAVDWITLSSPAQGVSCVSKMPPQNPDLTQI 1009  
 Qy 1053 RRAPQLNHSDDVYSINCLRLVPNQ-EINPHLLGNL---WLRL---KALKYKSMKIMV 1105  
 Db 1010 QRRSVLDCSTADCLHFRCDIPSLDIQDELFIPLGNLSFGWVSCQTLQEKVLLVSEAITF 1069  
 Qy 1106 NAALQRFHSPFIFREEDPSQIEFEISKOEDMOV--PIMLVGSTLGGILLALALVAL 1163  
 Db 1070 DTSVYSLQPGQEAFLR---AQVETTL---EYVYVYPIFLVAGSSVGGILLALITVL 1122  
 Qy 1164 RKLGFTRSARRRREPGLD 1181  
 Db 1123 YKLGFE---XKQYKEMLD 1137

RESULT 13  
 US-09-350-259-55  
 ; Sequence 55, Application US/09350259  
 ; Patent No. 6620915  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gallatin, Michael W.  
 ; APPLICANT: Van der Vieren, Monica  
 ; TITLE OF INVENTION: NO. 8620915el Human 2  
 ; FILE REFERENCE: 27866/35004  
 ; CURRENT APPLICATION NUMBER: US/09/350,259  
 ; CURRENT FILING DATE: 1999-07-08  
 ; EARLIER APPLICATION NUMBER: 09/193,043  
 ; EARLIER FILING DATE: 1998-11-16  
 ; EARLIER APPLICATION NUMBER: 08/173,497  
 ; EARLIER FILING DATE: 1993-12-23  
 ; EARLIER APPLICATION NUMBER: 08/286,889  
 ; EARLIER FILING DATE: 1994-08-05  
 ; EARLIER APPLICATION NUMBER: 08/362,652  
 ; EARLIER FILING DATE: 1994-12-21  
 ; EARLIER APPLICATION NUMBER: 08/943,363  
 ; EARLIER FILING DATE: 1997-10-03  
 ; NUMBER OF SEQ ID NOS: 114  
 ; SOFTWARE: Patent in Ver. 2.0  
 ; SEQ ID NO 55  
 ; LENGTH: 1161  
 ; TYPE: PRT  
 ; ORGANISM: Rattus rattus  
 US-09-350-259-55

Query Match 19.0%; Score 1184.5; DB 4; Length 1161;

Best Local Similarity 29.6%; Pred. No. 1,8e-97;  
 Matches 361; Conservative 207; Mismatches 519; Indels 131; Gaps 43;  
 Qy 7 LVVAVALSMPGFTDTFNDRKPRVIPSRTAFPGYTVQOHDIGNKWLVGABLETNG 66  
 Db 8 LLCGVFLASCHG---SMLDVEEP-IVFREDAAFGQTVQF---GGSELVVGAFLA 59  
 Qy 67 YQKTGDVYKCPVIHGNCTKLNLRVTLNSVSRKKNMRLGLSLATNPKNSFLACSPWS 126  
 Db 60 VNQTRLVDCAATGWCQPIVL-RSPLAV-----NMSLGLSLVATATNNAQLLACQPTAQ 113  
 Qy 127 HECGSSYYTTGCMRSVNSNFRFSKTVAPALQRC-QTYMDIVIVLSDGNSI---YPWVEVOH 183  
 Db 114 RACVKNMVAKSGCLLGGSLQIQAVPASMPCEPRQEMDIAFLIDGSGINORDFAQMKD 173  
 Qy 184 FLINILKFPYIGPGQIQCVVQYQEDVWHEPHLNDYRSVKVWVBAASHTEQGGTETRTA 243  
 Db 174 FYKALMGF--ASTSTLSLMQISNLKTHFTTTEFNILDPQSLVDPIVQLQGL-TYTA 230  
 Qy 244 FGIEFARSEAF--QKRGKRAKVMVITDGSNDSP-DLEKVIQOESERNVTRVAVAVL 300  
 Db 231 TGIPTVMEELFHSKNGSRKSKILLVITDQKQRPDLEYSVDVIPAADKAGIRVAVGV 290  
 Qy 301 GYNNRGINPETFLEIKVLIASDDPKHFNVTDEAALKDIVDALGDRIFSLSGT-NKHE 359  
 Db 291 DAFQE-----PTALKELNTIGSAPPQDHRVFKVGNFAULRSIORQOEKIFALEGTQSRSS 345  
 Qy 360 TSFGLMSQTSFSSHVVEDGVLGAVGAYDNWAGVAKTSACKVIP--LRSEYLKEFPPEE 417  
 Db 346 SSGFQHEMSQEGSSALTSQDGLVGVGVSFSSGSAF-----LYPPTTRPTFNMQS 398  
 Qy 418 LKHGAYLYTYSVSSQGRVYVAGAPRPHITKVKILFTWHNNRSLTIHQMRGQOIG 477  
 Db 399 VDMRDSYLGYS-TAVAFMKGVHSLIIGAPRHQHTGKWIPT-OEAKHWKPKSEVRQTQIG 456  
 Qy 478 SYFGSITSDIDGQVTVLLVGAFFMYFNEGRKGVVYVELR--QNEFVYNGTLDKSH 535  
 Db 457 SYFGSLCSVDVDRGSDXDLVIGAPHYEQTR-GGOVSFVPPGVGRWQCEATLHGEQ 515  
 Qy 536 SYQNAFSGSIASVRDLNODSYNDVVVGAPELNDHAGAIYIFHG-FRGSILKTPKRIIA 594  
 Db 516 GHFWGRFGVALTVLGDVNGDNLADVAIGAPEGSEESGAVYIPEGASRLRIMPSPSRVTG 575  
 Qy 595 SELATCLOYGSGSIHQDLNEDGLDLAVGALGNVILWSRFPVAINASLPEPSKINI 654  
 Db 576 SLSLRLQYFGQSLGGQDITQDGLVDLAVGAQHVLRLSLPLKVKVLSIRPAPMEVAK 635  
 Qy 655 FHRDCRSGRDATCL---AAFLCFTPIFLAPHFQTTTVG-IRYNATMDERRYTPRAHLDE 710  
 Db 636 AVQCWE--RTPTVLEAGEATVCLTVHKGSFDLLGNVQGSVRYDLALDPCRLISRIFDE 693  
 Qy 711 GDRFTNRVAVLLSSGQELCERINFHVL-DADYVVKVTFVSVEYL-----EDPDHGMPLD 764  
 Db 694 TKNCTLTGRKTLGLGDH-CETVKLLLPDCEVDAVSPILRLNLSLRVDSASPRNLHPVLA 752  
 Qy 765 DGNPTTLRVSVFPMWNGCNDEHCVDPDLVDARSDELPTAMEYCORVLRKPAQDCSAYTLSF 824  
 Db 753 VGSQDHIATSLPFKCKQELLCEGDL-----GISFNF 785  
 Qy 825 DTTVTIESTRORVAVATLENGENAYSTVLNISQSANLQFASLI---QKEDSDGSIEC 881  
 Db 786 SGLQVLVVGSPSLTIVTVWNEGEDSYGLTVFYPYAGLSYRVTGTQOPHYPLRLAC 845  
 Qy 882 VNE---ERRLQKQVNCVSYFFRAKAKAVAFRLDSEFSKIFL-HHLEIELAAGSDSNERD 937  
 Db 846 EAEPAQEDLRSSSCSINHPIFREGAKTTFMIFDVSKAFGLDRLLRKAKASSENKPD 905  
 Qy 938 STKEDNVAPLRFHLKYRADVLFTRSSLSHYEVLKNSLSRYDGIQPPFSCIPIRQNLGL 997  
 Db 906 TNK--TAFQLELPVKYTVYTLISRQEDSTNH-VNFSS---HGRROEAAHRYRVNLS 959  
 Qy 998 FEIHHMMKITIPIATRSQNLRLKLRDLFTDANTSCNIWGNSTERYPTVEE-----DL 1052

Db 960 LKL-AVRNFWPVL-----LNGVAVDVTLSSPAQGVSVSQMKPQNPDLTQI 1009  
QY 1053 RRAPOLNHSNVSINCNIRLVPNQ-EINEHLLGNL---WRLSL---KALKYKSMKIMV 1105  
Db 1010 QRRSLDCSIADCLFRCDIPSLD:QBELDFILRGSLFGWVSQTLQKVLVSEAITP 1069  
QY 1106 NAALQORQSPRIFREEDPSRQIEPEISKQEDMOV--PIWIVGSTLGLLALLLVLAL 1163  
Db 1070 DTSVYSQLEGGAEFLR---AQVEITL---BEVVYVEIFLAVAGSVGGLLLITVVL 1122  
QY 1164 RKLGFRRSARRRREGLD 1181  
Db 1123 YKLGf---XKROYKEMLD 1137

## RESULT 14

US-08-286-889-46  
; Sequence 46, Application US/08286889  
; Patent No. 5470953  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Mich  
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: P38,659  
; REFERENCE/DOCKET NUMBER: 27866/32168  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1155 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-286-889-46

Query Match 19.0%; Score 1181.5; DB 1; Length 1155;  
Best Local Similarity 30.0%; Pred. No. 3.4e-97;  
Matches 373; Conservative 199; Mismatches 501; Indels 171; Gaps 50;  
QY 5 RGLVW---AWALSLWPGFTDFTMDTRKPRVPGSRTAFPGYTVQHQHDSGNKMLVVGAP 61  
Db 3 RGVVILLCWALASCHG-----SNLDVEKP-VYFKEDASFGQTVVQF---GGSRLVVGAP 54  
QY 62 LETNGYQKTDGVYKCPVHNGCTKLNLRGVTLSNVSERKDNRLGLSLATNPKNSFLAC 121  
Db 55 LEAVAVNQGSQSDCPAPGVCQPIILL-HIPLEAV-----NMSIGLSLVADTNNSQLEAC 108  
QY 122 SPLWSEHCSSYTTGMCNRSVNSNFRFSKTVAPALQRCQ-TYMDIVIVLDGNSNI--VPW 178

Db 109 GPTAQRACAKMYAKGSCLLLGSSLSQFIQAIPATMPECPCQEMDIAFLDGSQSIDOSDF 168  
QY 179 VEVQHFLINILKXPIYGPQI-----QGVVVQYGEDVWHFHLNDYRSVKDVVAAASHIE 233  
Db 169 TOMKDFVKALM-----GQLASTSTSFSLAQYSNILKTHFTPTTFKSSLSQSLVDAIV 221  
QY 234 QRGQTEHTAFGIEPARSEAF--CKGGRKGAKKVIWITDGESHUSP-DLEKVIQOQSERD 290  
Db 222 QLQGL-TYTAGCIQKVVKELPHSKNGARKSAKILIVITDQKFRDPLEYHVLPEAKA 280  
QY 291 NVTRYAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDVLDAIGDRIF 350  
Db 281 GIIRYALGVGDFAFE-----PTALQELNTICSPSQDHVFKVGNFVALRSIORIQEKIF 335  
QY 351 SLECF-NKNETSFCLEMSOTGSSHVVVEDGVLLGAVGAYNGAVLAKETSAGKVIPLRES 409  
Db 336 AIEGTERSSSSPOHEMSQEGFSALSNDGFLVCAVCGFSWGAFLYPS-----NKRST 390  
QY 410 YLKEFFPEELKNHGAYLCYTVTSVSSRQGVYVAGAPRNNHTKVLFTMNNRSLTHQ 469  
Db 391 FINMSQENEDMRDAYLGYS-TALAFWKGVHSLILGAPRHQHTGKVIPT-QESRHRPEKS 448  
QY 470 AMRQOIGSYFGSEITSDVIDGQVTDVLLVGAEMYNEGRERKVVYVELR--QNRVY 527  
Db 449 EVRGTOIGSYFGASLCSVDMDRDGSITDLVIGVPHYEHTR-GGQSVCPMPGVRSRWHC 507  
QY 528 NGTLKDSHSYQNAFSGSIIASVRDLNQDSYNDVYVVGAPLEBNDHAGAIYIPHG-FRGSILK 586  
Db 508 GTTLHGEGCHPWGRFGAALTVLGDVNGDSLADVAIGAPEGBEENRGAVYIPHGASRQDLAP 567  
QY 587 TPQRITASELATGLQYFGCSIHQGLDNEGLDLAVALGALNAVILMSRVVQVAINSLH 646  
Db 568 SPSQRTVGSQFLRLQYFGQSLGGQDLTQGLDVLAVAGQGHVLLRLSLPLKVGISIR 627  
QY 647 PEPSKINIFHRCKRSGRDATCL---AALFLCFT-----PIPLAPHFQTTTIGIRYNATMD 698  
Db 628 PAPSEVAKTVYQC--WGRTPTVLEAGEATVCLTVRKSGSPDLG---DVQSSVAYDLALD 681  
QY 699 ERYTPRAHLDEGGDRPTNRAVLLSSGOELCERINFHVL-DAYVKKPVTSEYSL---754  
Db 682 PGRILISRAIFDETRKNTLTKRKTGLGDH-CETHKLLLPDCVEDAVTPIILRLSLAGD 740  
QY 755 EDPDHG--PMLDDGWPTTLRVSVFPWNGCNEDECPDVLVDARSDLPTAMEYQORVLK 812  
Db 741 SAPSENLRPVLAVGSDHVTASPPF-----KXNCEGNL-----773  
QY 813 PAQDCSAYTLGSDTTVFIIBSTRQVAVEATLENGENAYSTVLNISQSANLQFASLIQK 872  
Db 774 -----GVSNFSGQLQVLEVGSSDELTVTVTWNEGEDSYGLIKFYYPABLSYRVTRA 827  
QY 873 ESDSG---SIECVNE---ERRLOQVCNVSVYPERAKAKVAFRLDSSEKSIPL-HHLEI 925  
Db 828 QQPHYPILRLACEABPTGQBSLRSSCSINHPIREGAKATFMITFDVSKAFGLDRLLL 887  
QY 926 ELAAGSDSNERDSTKBDNVAPLRFHLKYEADVLPFRSS-----SLSHYEVKLNLSLE 977  
Db 888 RASASSENKPETSK--TAFQLELPVKYTVTVTVISRQEDSTKTFNFSSSHGE-RQKGAEH 944  
QY 978 RYDGIGPPFCIFRIQNLGLFPFHGMKKITIPATRSNRL--LKLRDLPTDEANTSCN 1035  
Db 945 RY-----RVNKLSPILL-AISVNFVWPVILL-NGVAVMDVTLR---SPAQGVSC-987  
QY 1036 IWNSTBYRPTFVREEDURRAPO-----LNHSNSDVVSINCNIRLVPNQEIFHLLGNL--1088  
Db 988 -----VQSRBPQHSDDLTLTQIGESVLDCAIDCLHLRCDIPSLGTLDELDFILKGNLSF 1042  
QY 1089 -WRLSL---KALKYKSMKIMYNAALQRFHSPTIFREEDPSRQIEPEISKQEDQVWIWI 1144  
Db 1043 GWSISQTLQKVKLLSEAEITFNTSVYSQLPQGEAFRAQVSTMLEEYVYVE-----PVFL 1097  
QY 1145 IVGSTLGLLALLLVLALRKLGFRRSARRRE---PGLDPTP 1184  
Db 1098 WPFSSVGGLELLALITVALYKLGFFK--RQYKEMLDLPSADPDP 1139

## RESULT 15

US-08-485-618-46

; Sequence 46, Application US/08485618

; Patent No. 5728533

; GENERAL INFORMATION:

; APPLICANT: Gallatin, W. Michael

; APPLICANT: Van der Vliet, Monica

; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit

; NUMBER OF SEQUENCES: 103

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &amp; Borun

; STREET: 233 South Wacker Drive, 6300 Sear Tower

; City: Chicago

; STATE: Illinois

; COUNTRY: United States

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/485,618

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/173,497

; FILING DATE: 23-DEC-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/286,889

; FILING DATE: 5-AUG-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/362,652

; FILING DATE: 21-DEC-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Williams Jr., Joseph A.

; REGISTRATION NUMBER: 38,659

; REFERENCE/DOCKET NUMBER: 27866/32797

; TELEPHONE: 312-474-6300

; TELEFAX: 312-474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 46:

; TYPE: amino acid

; LENGTH: 1155 amino acids

; SEQUENCE CHARACTERISTICS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-485-618-46

Query Match 19.0%; Score 1181.5; DB 1; Length 1155;

Best Local Similarity 30.0%; Pred. No. 3.4e-97;

Matches 373; Conservative 199; Mismatches 501; Indels 171; Gaps 50;

5 RGLV---AWALSLPFGTDTFNMOTKRPVPGSRTPAFGTVYQOHDISGNKWLTVGAP 61

3 RGVVILLCGVALASCHG---SNLDVEKP-VVFKEDAAASFGQTVVQF---GGSLVVGAP 54

62 LETNGYQKTDVYKCPVIHGNCTKMLGRTVLSNVSRKONMRLGLSLATNPDKNSFLAC 121

55 LEAVAVNQGGSDCPATGVCQPILL-HIPLEAV-----NMSLGLSLVADTNNSQLLAC 108

122 SPLMSHECGSSYYTTCMCVRVNSNFRFSKTVAPALQRCQ-TYMDIVIVLDGNSI--YFW 178

109 GPTAQACAKNMYAKGSCILGSSLQFIQAIATMPCFCPGQEMDIAFLIDGSGSIDQSD 168

179 VEVOHFLINLKKFYTGPOI-----QGVWQGEDVVEHFNLDYRSVQVWVEAASHIE 233

169 TQKDFVFKALM-----GQLASTSTFSLMQVSNILKTHFTTFEKSLSQSLVDAIV 221

234 QRGGTETRTAFGIEPARSEAF--QKGRGKGAKMIVITDGSBHDSP-DLEKVIQDSRD 290

Search completed: June 24, 2004, 17:56:46

Job time : 28 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 17:58:06 ; Search time 27 Seconds  
(without alignments)  
4232.427 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 1188

Sequence: 1 MDLPRGLVVAWLSLWPGFT.....PFSARRRRRGDLTPFKVLE 1188

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : PIR 78:\*\*

1: pir1:\*\*

2: pir2:\*\*

3: pir3:\*\*

4: pir4:\*\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 15    | 1.3         | 272    | 2 A55348 | integrin alpha-1 - |
| 2          | 15    | 1.3         | 1151   | 2 A45226 | integrin alpha-1 c |
| 3          | 15    | 1.3         | 1180   | 2 A35854 | integrin alpha-1 c |
| 4          | 11    | 0.9         | 1170   | 2 I45914 | integrin alpha-2 s |
| 5          | 11    | 0.9         | 1181   | 2 A33998 | integrin alpha-2 s |
| 6          | 9     | 0.8         | 74     | 2 I51524 | integrin alpha-2 s |
| 7          | 9     | 0.8         | 315    | 2 A35567 | permease (imported |
| 8          | 9     | 0.8         | 371    | 1 T43407 | 3-isopropylmalate  |
| 9          | 9     | 0.8         | 607    | 2 S60658 | legumin - Gnetum g |
| 10         | 9     | 0.8         | 1178   | 2 S44142 | VLA-2 protein homo |
| 11         | 8     | 0.7         | 20     | 2 A60822 | cytochrome P450 PB |
| 12         | 8     | 0.7         | 42     | 2 A34259 | cytochrome P450mt4 |
| 13         | 8     | 0.7         | 76     | 2 I51527 | integrin alpha 5 s |
| 14         | 8     | 0.7         | 76     | 2 A45337 | heat-stable antige |
| 15         | 8     | 0.7         | 76     | 2 I53107 | CD24 precursor - r |
| 16         | 8     | 0.7         | 80     | 2 A40856 | S cell surface ant |
| 17         | 8     | 0.7         | 103    | 2 A40856 | conserved hypother |
| 18         | 8     | 0.7         | 127    | 2 B75301 | hypothetical prote |
| 19         | 8     | 0.7         | 141    | 2 I51785 | heat-stable antige |
| 20         | 8     | 0.7         | 224    | 2 E71228 | hypothetical prote |
| 21         | 8     | 0.7         | 226    | 2 S78800 | hypothetical prote |
| 22         | 8     | 0.7         | 234    | 1 S15102 | eosinophil major b |
| 23         | 8     | 0.7         | 264    | 2 S22090 | catechol O-methyl  |
| 24         | 8     | 0.7         | 308    | 2 AG2637 | conserved hypother |
| 25         | 8     | 0.7         | 356    | 2 P97419 | BH1459 conserved h |
| 26         | 8     | 0.7         | 360    | 2 A85016 | hypothetical prote |
| 27         | 8     | 0.7         | 367    | 2 S19172 | cytochrome P450 2B |
| 28         | 8     | 0.7         | 387    | 1 DYH04  | dopamine receptor  |
| 29         | 8     | 0.7         | 402    | 1 S23860 | chloramphenicol re |

|     |   |     |      |          |                     |
|-----|---|-----|------|----------|---------------------|
| 30  | 8 | 0.7 | 415  | 2 G83568 | probable permeal    |
| 31  | 8 | 0.7 | 442  | 2 AG3504 | dihydrofolate synt  |
| 32  | 8 | 0.7 | 444  | 2 T01721 | hypothetical prote  |
| 33  | 8 | 0.7 | 487  | 2 T47107 | benzaldehyde dehyd  |
| 34  | 8 | 0.7 | 491  | 1 O4RTPB | cytochrome P450 2B  |
| 35  | 8 | 0.7 | 491  | 1 O4RTP2 | cytochrome P450 2B  |
| 36  | 8 | 0.7 | 500  | 2 B31047 | testosterone 16alp  |
| 37  | 8 | 0.7 | 547  | 2 A31314 | maolactac enzyme    |
| 38  | 8 | 0.7 | 547  | 2 A31686 | maolactac enzyme    |
| 39  | 8 | 0.7 | 605  | 1 W1WLEB | E1 protein - bovin  |
| 40  | 8 | 0.7 | 606  | 2 A72429 | oligopeptide ABC t  |
| 41  | 8 | 0.7 | 614  | 2 A59845 | Na+/H+ antiporter   |
| 42  | 8 | 0.7 | 620  | 1 W1WLB2 | E1 protein - bovin  |
| 43  | 8 | 0.7 | 685  | 2 AC0527 | ferriochrome transp |
| 44  | 8 | 0.7 | 697  | 2 T16306 | hypothetical prote  |
| 45  | 8 | 0.7 | 739  | 2 A88445 | protein C26E6.10 l  |
| 46  | 8 | 0.7 | 849  | 2 B83349 | probable C1pA/B-t   |
| 47  | 8 | 0.7 | 1041 | 2 T31437 | integrin alpha cha  |
| 48  | 8 | 0.7 | 1065 | 2 T25068 | hypothetical prote  |
| 49  | 8 | 0.7 | 1146 | 2 S40311 | integrin - fruit f  |
| 50  | 8 | 0.7 | 1153 | 1 RWHUB  | cell surface glyco  |
| 51  | 8 | 0.7 | 1163 | 1 RWHUB  | cell surface glyco  |
| 52  | 8 | 0.7 | 1179 | 2 A53213 | integrin alpha-B c  |
| 53  | 8 | 0.7 | 1366 | 2 S77664 | IGA-specific metal  |
| 54  | 8 | 0.7 | 1460 | 2 D81675 | polymorphic membra  |
| 55  | 8 | 0.7 | 1664 | 2 S67250 | DNA-directed RNA p  |
| 56  | 8 | 0.7 | 1802 | 2 H88444 | protein C26E6.12 l  |
| 57  | 8 | 0.7 | 2238 | 1 RRVUBY | genome polypotein   |
| 58  | 7 | 0.6 | 47   | 2 D81686 | hypothetical prote  |
| 59  | 7 | 0.6 | 74   | 2 T47376 | hypothetical prote  |
| 60  | 7 | 0.6 | 89   | 2 I49515 | B144 protein B - m  |
| 61  | 7 | 0.6 | 94   | 2 A99774 | probable excisiona  |
| 62  | 7 | 0.6 | 99   | 2 P85636 | probable excisiona  |
| 63  | 7 | 0.6 | 103  | 2 G75257 | hypothetical prote  |
| 64  | 7 | 0.6 | 104  | 2 PH0141 | integrin alpha cha  |
| 65  | 7 | 0.6 | 106  | 4 S57386 | hypothetical prote  |
| 66  | 7 | 0.6 | 107  | 2 JC1127 | major allergen cha  |
| 67  | 7 | 0.6 | 109  | 2 C56413 | insulin precursor   |
| 68  | 7 | 0.6 | 110  | 1 INRB   | insulin precursor   |
| 69  | 7 | 0.6 | 110  | 2 S30268 | protein hdea precu  |
| 70  | 7 | 0.6 | 110  | 2 S30268 | protein hdea precu  |
| 71  | 7 | 0.6 | 110  | 2 G86023 | protein hdea precu  |
| 72  | 7 | 0.6 | 110  | 2 G86023 | hypothetical prote  |
| 73  | 7 | 0.6 | 114  | 2 C95870 | hypothetical prote  |
| 74  | 7 | 0.6 | 115  | 2 D75434 | conserved hypother  |
| 75  | 7 | 0.6 | 132  | 2 B81215 | hypothetical prote  |
| 76  | 7 | 0.6 | 133  | 2 H75386 | hypothetical prote  |
| 77  | 7 | 0.6 | 139  | 1 C45051 | lamprin 2 precursor |
| 78  | 7 | 0.6 | 139  | 2 A38612 | insulin-like growt  |
| 79  | 7 | 0.6 | 141  | 1 HACH1  | hemoglobin alpha-D  |
| 80  | 7 | 0.6 | 141  | 2 S56103 | hemoglobin alpha-D  |
| 81  | 7 | 0.6 | 142  | 2 JT0573 | retinoic acid-indu  |
| 82  | 7 | 0.6 | 142  | 2 G87647 | hypothetical prote  |
| 83  | 7 | 0.6 | 143  | 2 T05564 | hypothetical prote  |
| 84  | 7 | 0.6 | 154  | 2 S14947 | 2S albumin - Brazi  |
| 85  | 7 | 0.6 | 154  | 2 P83074 | conserved hypother  |
| 86  | 7 | 0.6 | 158  | 2 G82494 | conserved hypother  |
| 87  | 7 | 0.6 | 160  | 1 P70351 | NADH2 dehydrogenas  |
| 88  | 7 | 0.6 | 160  | 2 S58214 | apoptosis inducer   |
| 89  | 7 | 0.6 | 161  | 2 D71903 | hit family protein  |
| 90  | 7 | 0.6 | 163  | 2 S66795 | probable membrane   |
| 91  | 7 | 0.6 | 163  | 2 E72225 | conserved hypother  |
| 92  | 7 | 0.6 | 171  | 1 B64498 | conserved hypother  |
| 93  | 7 | 0.6 | 176  | 2 AC0207 | probable exported   |
| 94  | 7 | 0.6 | 176  | 2 T38925 | cytochrome oxidase  |
| 95  | 7 | 0.6 | 176  | 2 T46709 | hypothetical prote  |
| 96  | 7 | 0.6 | 176  | 2 A81250 | hypothetical prote  |
| 97  | 7 | 0.6 | 179  | 2 A64551 | hypothetical prote  |
| 98  | 7 | 0.6 | 179  | 2 D71957 | hypothetical prote  |
| 99  | 7 | 0.6 | 180  | 2 E70358 | HUPE hydroxynase r  |
| 100 | 7 | 0.6 | 182  | 2 E82367 | hypothetical prote  |

A35854  
integrin alpha-1 chain precursor - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 23-Oct-1990 #sequence\_revision 13-Sep-1991 #text\_change 15-Sep-2003  
C;Accession: A35854; S11243  
R;Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J. Cell Biol. 111, 709-720, 1990  
A;Title: Molecular cloning of the rat integrin alpha-1-subunit: a receptor for laminin A  
A;Reference number: A35854; MUID:90338125; PMID:2380249  
A;Accession: A35854  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1180 <IGN>  
A;Cross-references: GB:X52140; NID:g56493; PIDN:CAA36384.1; PID:g56494  
C;Keywords: cell adhesion; cytoskeleton; transmembrane protein  
F;170-345/Domain: von Willebrand factor type A repeat homology <VWA2>

2. TTB position: 5q11.1-5q11.2

## RESULT 8

Query Match 0.8%; Score 9; DB 2; Length 607;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 9; Conservative 0; Mismatches 0; Indels

QY 1153 LLLALLLV 1161  
 |||||  
 Db 10 LLLALLLV 18

## RESULT 10

S44142  
 VLA-2 protein homolog - mouse  
 A:Species: Mus musculus (house mouse)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 15-Sep-2003  
 C:Accession: S44142  
 R:Delman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze  
 submitted to the EMBL Data Library, January 1994  
 A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not  
 A:Reference number: S44142  
 A:Accession: S44142  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1178 <EDE>  
 A:Cross-references: EMBL:Z29987; NID:G473099; PIDN:CAA82877.1; PID:G473099  
 F:169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 0.8%; Score 9; DB 2; Length 1178;  
 Best Local Similarity 100.0%; Pred. No. 7.4; Mismatches 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0

QY 495 TDVLLVGAP 503  
 |||||  
 Db 503 TDVLLVGAP 511

## RESULT 11

A60822  
 cytochrome P450 PB-3a - rat (fragment)  
 N:Contains: oxidoreductase (EC 1.-.-.-)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 11-Jun-1999  
 C:Accession: A60822; I55191  
 R:Ameliaz, Z.; Narbonne, J.F.; Wolf, C.R.; Robertson, L.W.; Oesch, F.  
 Biochem. Pharmacol. 37, 3245-3249, 1988  
 A:Title: Effect of nutritional imbalances on cytochrome P-450 isozymes in rat liver.  
 A:Reference number: A60822; MUID:88293549; PMID:3041969  
 A:Accession: A60822  
 A:Status: preliminary  
 A:Molecule type: protein

A:Residues: 1-20 <AME>  
 A:Keywords: electron transfer; heme; monooxygenase; oxidoreductase; transmembrane protein

Query Match 0.7%; Score 8; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 2.2; Mismatches 0; Indels 0; Gaps 0;  
 Matches 8; Conservative 0

QY 1153 LLLALLLV 1160  
 |||||  
 Db 6 LLLALLLV 13

## RESULT 12

A34259  
 cytochrome P450mt4, phenobarbital-inducible, mitochondrial, hepatic - rat (fragment)  
 N:Contains: oxidoreductase (EC 1.-.-.-)  
 C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 05-Mar-1999  
 C:Accession: A34259  
 R:Shayiq, R.M.; Avadhani, N.G.  
 Biochemistry 29, 866-873, 1990  
 A:Title: A phenobarbital-inducible hepatic mitochondrial cytochrome P-450 immunochemi  
 A:Reference number: A34259; MUID:90254127; PMID:2340279  
 A:Accession: A34259  
 A:Molecule type: protein  
 A:Residues: 1-42 <SHA>  
 A:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
 C:Keywords: electron transfer; heme; liver; mitochondrion; monooxygenase; oxidoreduct

Query Match 0.7%; Score 8; DB 2; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 4.1; Mismatches 0; Indels 0; Gaps 0;  
 Matches 8; Conservative 0

QY 1153 LLLALLLV 1160  
 |||||  
 Db 6 LLLALLLV 13

## RESULT 13

I51527  
 integrin alpha 5 subunit - African clawed frog (fragment)  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 29-Sep-1999  
 C:Accession: I51527  
 R:Whittaker, C.A.; DeSimone, D.W.  
 Development 117, 1239-1249, 1993  
 A:Title: Integrin alpha subunit mRNAs are differentially expressed in early Xenopus e  
 A:Reference number: I51524; MUID:94008528; PMID:8404528  
 A:Accession: I51527  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-76 <WHI>  
 A:Cross-references: GB:L10191; NID:G214544; PIDN:AAA16249.1; PID:G214545  
 C:Superfamily: integrin alpha-2b chain

Query Match 0.7%; Score 8; DB 2; Length 76;  
 Best Local Similarity 100.0%; Pred. No. 6.9; Mismatches 0; Indels 0; Gaps 0;  
 Matches 8; Conservative 0

QY 541 RFGSSIAS 548  
 |||||  
 Db 63 RFGSSIAS 70

## RESULT 14

A43537  
 heat-stable antigen M1/69-J1ld precursor - mouse  
 N:Alternate names: CD24 protein; nectadrin  
 C:Species: Mus musculus (house mouse)  
 C:Date: 06-Nov-1992 #sequence\_revision 06-Nov-1992 #text\_change 31-Jan-2000  
 C:Accession: A43537; I48287; S15784; S15783; S43709; S32240; S33129  
 R:Kay, R.; Takai, F.; Humphries, R.K.  
 J. Immunol. 145, 1952-1959, 1990  
 A:Title: Expression cloning of a cDNA encoding M1/69-J1ld heat-stable antigens.  
 A:Reference number: A43537; MUID:90361906; PMID:2118158  
 A:Accession: A43537  
 A:Molecule type: mRNA  
 A:Residues: 1-76 <KAY>  
 A:Cross-references: GB:M58661; NID:G198985; PIDN:AAA39481.1; PID:G198986  
 R:Wenger, R.H.; Rochelle, J.M.; Seidin, M.F.; Kohler, G.; Nielsen, P.J.  
 J. Biol. Chem. 268, 23345-23352, 1993  
 A:Title: The heat stable antigen (mouse CD24) gene is differentially regulated but ha  
 A:Reference number: A48876; MUID:94043127; PMID:8226859  
 A:Accession: I48287  
 A:Status: translation not shown; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-76 <RES>  
 A:Cross-references: EMBL:X72910; NID:G296466; PIDN:CAA51415.1; PID:G296467  
 R:Wenger, R.H.; Ayane, M.; Bose, R.; Koehler, G.; Nielsen, P.J.  
 Eur. J. Immunol. 21, 1039-1046, 1991

A:Title: The genes for a mouse hematopoietic differentiation marker called the heat-stab  
A:Reference number: S15783; MUID:91209380; PMID:2019286  
A:Accession: S15784  
A:Molecule type: DNA  
A:Residues: 1-76 <WE3>  
A:Cross-references: EMBL:X56469; NID:G51439; PIDN:CAA39841.1; PID:G51440  
A:Accession: S15783  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 32-76 <WE2>  
A:Cross-references: EMBL:X53825  
R.Nielsen, P.J.  
submitted to the EMBL Data Library, July 1990  
A:Reference number: S13111  
A:Accession: S43709  
A:Molecule type: mRNA  
A:Residues: 1-76 <NIE>  
A:Cross-references: EMBL:X53825; NID:G55441; PIDN:CAA37822.1; PID:G55442  
C:Genetics:  
A:Gene: Cd24a  
A:Map position: 10  
A:Introns: 23/3  
A:Keywords: blocked carboxyl end; glycoprotein; lipoprotein; membrane protein; phosphatidyl  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:27-56/Product: heat-stable antigen ML/69-Jiid #status predicted <MAT>  
F:57-76/Domain: carboxyl-terminal propeptide #status predicted <Crp>  
F:57/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Ser) (in mature form)

Query Match 0.7%; Score 8; DB 2; Length 76;  
Best Local Similarity 100.0%; Pred. No. 6.9; Mismatches 0; Indels 0; Gaps 0;  
Matches 8; Conservative 0

QY 1152 GLLLLALL 1159  
|||||  
Db 12 GLLLLALL 19

RESULT 15  
IS3107  
CD24 precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999  
C:Accession: IS3107; S25146  
R.Shirasawa, T.; Akashi, T.; Sakamoto, K.; Takahashi, H.; Maruyama, N.; Hirokawa, K.  
Dev. Dyn. 198, 1-13, 1993  
A:Title: Gene expression of CD24 core peptide molecule in developing brain and developin  
A:Reference number: IS3107; MUID:94122434; PMID:8292828  
A:Accession: IS3107  
A>Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-76 <RES>  
A:Cross-references: EMBL:Z11663; NID:G55901; PIDN:CAA77731.1; PID:G55902  
C:Keywords: phosphatidylinositol linkage

Query Match 0.7%; Score 8; DB 2; Length 76;  
Best Local Similarity 100.0%; Pred. No. 6.9; Mismatches 0; Indels 0; Gaps 0;  
Matches 8; Conservative 0

QY 1152 GLLLLALL 1159  
|||||  
Db 12 GLLLLALL 19

RESULT 16  
A48996  
B cell surface antigen CD24 precursor - human  
N:Alternate names: cluster-w4 antigen; signal transducer CD24  
C:Species: Homo sapiens (man)  
C:Date: 13-Dec-1993 #sequence\_revision 02-Aug-1996 #text\_change 21-Jul-2000  
C:Accession: IS6114; A48996; I54201; I72676  
R.Kay, R.; Rosten, P.M.; Humphries, R.K.  
J. Immunol. 147, 1412-1415, 1991  
A:Title: CD24, a signal transducer modulating B cell activation responses, is a very shd

A:Reference number: IS6114; MUID:91332458; PMID:1831224  
A:Accession: IS6114  
A>Status: translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-80 <RES>  
A:Cross-references: GB:L33930; NID:G500848; PID:G500849; GB:M58664; NID:G180167; PID:R.Jackson, D.; Walbel, R.; Weber, E.; Bell, J.; Stahel, R.A.  
Cancer Res. 52, 5264-5270, 1992  
A:Title: CD24, a signal-transducing molecule expressed on human B cells, is a major s  
A:Reference number: A48996; MUID:93007871; PMID:1327504  
A:Accession: A48996  
A:Molecule type: mRNA  
A:Residues: 1-56, 'V', '58-80 <JAC>  
A:Cross-references: GB:X69397; GB:S44888; NID:G996167; PIDN:CAA49195.1; PID:G996168  
A:Experimental source: small cell lung carcinoma line SW2  
A>Note: sequence extracted from NCBI backbone (NCBIN:114635, NCBIF:114636)  
A>Note: both 57-Val and 57-Ala were found in small cell carcinoma line DC571  
R.Hough, M.R.; Rosten, P.M.; Sexton, T.L.; Kay, R.; Humphries, R.K.  
Genomics 22, 154-161, 1994  
A:Title: Mapping of CD24 and homologous sequences to multiple chromosomal loci.  
A:Reference number: I54201; MUID:95048364; PMID:7959762  
A:Accession: I54201  
A>Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 5-11, 'W', '13-43, 'T', '45-80 <RE2>  
A:Cross-references: GB:S75311; NID:G83385; PIDN:AAD14170.1; PID:G4261870  
C:Genetics:  
A:Gene: GDB:CD24  
A:Cross-references: GDB:383795; OMIM:600074  
A:Map position: 6q21-6q21  
C:Keywords: B-cell; blocked carboxyl end; glycoprotein; lipoprotein; phosphatidylinos  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-67/Product: B cell surface antigen CD24 #status predicted <MAT>  
F:68-80/Domain: carboxyl-terminal propeptide #status predicted <CPRO>  
F:67/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Ser) (in mature fo

Query Match 0.7%; Score 8; DB 2; Length 80;  
Best Local Similarity 100.0%; Pred. No. 7.3; Mismatches 0; Indels 0; Gaps 0;  
Matches 8; Conservative 0

QY 1152 GLLLLALL 1159  
|||||  
Db 12 GLLLLALL 19

RESULT 17  
AF0856  
conserved hypothetical protein STY3056 [imported] - Salmonella enterica subsp. enteri  
C:Species: Salmonella enterica subsp. enterica serovar Typhi  
A>Note: this species has also been called Salmonella typhi  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C:Accession: AF0856  
R.Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church  
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AF0856  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-103 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD06037.1; PID:G16504004; GSPDB:GN00176  
C:Genetics:  
A:Gene: STY3056  
C:Superfamily: hypothetical protein HI0673

Query Match 0.7%; Score 8; DB 2; Length 103;  
Best Local Similarity 100.0%; Pred. No. 9; Mismatches 0; Indels 0; Gaps 0;  
Matches 8; Conservative 0

QY 1153 LLLLLALL 1160

```

Db      6 LLLLLLLV 13
|||||
RESULT 18
B75301
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C:Accession: B75301
R:White, O.; Eisen, J.A.; Heidelberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: B75301
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-127 <WHI>
A:Cross-references: GB:AE002054; GB:AE000513; NID:g6460010; PIDN:AAF11763.1; PID:g646001
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2210
A:Map position: 1
C:Superfamily: Deinococcus radiodurans hypothetical protein DR2210
Query Match      0.7%  Score 8;  DB 2;  Length 127;
Best Local Similarity 100.0%;  Pred. No. 11;
Matches      8;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1152 GLLLALL 1159
|||||
Db      7 GLLLALL 14
|||||
RESULT 19
S15785
heat-stable antigen-related hypothetical protein HSA-C - mouse
C:Species: Mus musculus (house mouse)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C:Accession: S15785
R:Wenger, R.H.; Ayane, M.; Bose, R.; Koehler, G.; Nielsen, P.J.
Eur. J. Immunol. 21, 1039-1046, 1991
A:Title: The genes for a mouse hematopoietic differentiation marker called the heat-stab
A:Reference number: S15783; MUID:91209380; PMID:2019286
A:Accession: S15785
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-141 <WEN>
A:Cross-references: EMBL:X56486; NID:g51441; PIDN:CAA19843.1; PID:g51442
A:Note: the authors translated the codon TTC for residue 67 as Ser and AGA for residue 1
A:Note: the authors did not translate the codon for residue 101
C:Genetics:
A:Introns: #status absent
C:Keywords: glycoprotein; membrane protein; phosphatidylinositol linkage
Query Match      0.7%  Score 8;  DB 2;  Length 141;
Best Local Similarity 100.0%;  Pred. No. 12;
Matches      8;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1152 GLLLALL 1159
|||||
Db      12 GLLLALL 19
|||||
RESULT 20
E71228
hypothetical protein PH0087 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: E71228
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

```

```

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Ogi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: E71228
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-224 <KAW>
A:Cross-references: GB:AP000001; NID:g3236128; PIDN:BA239156.1; PID:g3256473
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBa
C:Genetics:
A:Gene: PH0087
Query Match      0.7%  Score 8;  DB 2;  Length 224;
Best Local Similarity 100.0%;  Pred. No. 18;
Matches      8;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      512 RGKVVVYE 519
|||||
Db      105 RGKVVVYE 112
|||||
RESULT 21
S76800
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S76800
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima,
O., K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76800
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-226 <KAN>
A:Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BAAL8712.1; PID:g165
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: conserved hypothetical protein MTH747
Query Match      0.7%  Score 8;  DB 2;  Length 226;
Best Local Similarity 100.0%;  Pred. No. 18;
Matches      8;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1073 RLVPNQEI 1080
|||||
Db      183 RLVPNQEI 190
|||||
RESULT 22
S15102
eosinophil major basic protein 2 precursor - guinea pig
C:Species: Cavia porcellus (guinea pig)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jul-2000
C:Accession: S15102; S18501
R:Aoki, I.; Shindoh, Y.; Nishida, T.; Nakai, S.; Hong, Y.M.; Mio, M.; Saito, T.; Tasa
FBS Lett. 282, 56-60, 1991
A:Title: Comparison of the amino acid and nucleotide sequences between human and two
A:Reference number: S15102; MUID:91224343; PMID:2026266
A:Accession: S15102
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-234 <AOK>
A:Cross-references: DDBJ:D00817; NID:g3135095; PIDN:BA000697.1; PID:g2202933
A:Accession: S18501
A:Molecule type: protein
C:Superfamily: eosinophil major basic protein precursor; C-type lectin homology
C:Keywords: antibiotic; chondroitin sulfate proteoglycan; cytotoxin; eosinophil; glyc

```

F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-115/Domain: propeptide #status predicted <PRO>  
F:116-234/Product: eosinophil major basic protein 2 #status experimental <MAT>  
F:117-232/Domain: C-type lectin homology <LCH>  
F:24,25/Binding site: carbohydrate (Ser) (covalent) #status predicted  
F:69/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted  
F:135-232,209-224/Disulfide bonds: #status predicted

Query Match 0.7%; Score 8; DB 1; Length 234;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160  
|||||||  
DB 4 LLLALLV 11

RESULT 23  
S22090  
catechol O-methyltransferase (EC 2.1.1.6) - rat  
N:Alternate names: catechol-O-methyltransferase  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 21-Jul-2000  
C:Accession: S22090; JQ0787; F52356  
R:Finhunen, J.; Salminen, M.; Jalanko, A.; Ukkonen, S.; Ulmanen, I.  
submitted to the EMBL Data Library, June 1992  
A:Description: Structure of the rat catechol-O-methyltransferase gene. Separate promoter  
A:Reference number: S22090  
A:Accession: S22090  
A:Molecule type: DNA  
A:Residues: 1-264 <FEN>  
A:Cross-references: EMBL:212651; NID:G55889; PIDN:CAA78276.1; PID:G55890  
R:Salminen, M.; Lundstroem, K.; Tilgmann, C.; Savolainen, R.; Kalkkinen, I.  
Gene 93, 241-247, 1990  
A:Title: Molecular cloning and characterization of rat liver catechol-O-methyltransferase  
A:Reference number: JQ0787; MUID:91033034; PMID:2227437  
A:Accession: JQ0787  
A:Molecule type: mRNA  
A:Residues: 44-264 <SAL>  
A:Cross-references: GB:M60753; NID:G203336; PIDN:AAA40881.1; PID:G203337  
A:Experimental source: liver  
R:Tenhunen, J.; Ulmanen, I.  
Biochem. J. 296, 595-600, 1993  
A:Title: Production of rat soluble and membrane-bound catechol O-methyltransferase forms  
A:Reference number: I52356; MUID:94107221; PMID:8280056  
A:Accession: I52356  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-264 <RES>  
A:Cross-references: EMBL:212651; NID:G55889; PIDN:CAA78276.1; PID:G55890  
C:Comment: This enzyme catalyzes the transfer of a methyl group from S-adenosyl-L-methio  
C:Genetics:  
A:Gene: COMT  
C:Superfamily: coffeoyl-CoA 3-O-methyltransferase  
C:Keywords: methyltransferase; S-adenosylmethionine; transmembrane protein

Query Match 0.7%; Score 8; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159  
|||||||  
DB 9 GLLLLALL 16

RESULT 24  
AG2637  
conserved hypothetical protein Atu0498 [imported] - Agrobacterium tumefaciens (strain C5  
C:Species: Agrobacterium tumefaciens  
C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 19-Nov-2002  
C:Accession: AG2637  
R:Wood, D.W.; Scubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell

: Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AG2637  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-308 <KUR>  
A:Cross-references: GB:AB008688; PIDN:AA441517.1; PID:G17738846; GSPDB:GN00186  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu0498  
A:Map position: circular chromosome

Query Match 0.7%; Score 8; DB 2; Length 308;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1155 LLALLVLA 1162  
|||||||  
DB 275 LLALLVLA 282

RESULT 25  
F97419  
BH1459 conserved hypothetical protein (AP001512) [imported] - Agrobacterium tumefacie  
C:Species: Agrobacterium tumefaciens  
C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
C:Accession: F97419  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldm  
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium  
A:Reference number: A97359; MUID:21608551; PMID:11743194  
A:Accession: F97419  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-356 <KUR>  
A:Cross-references: GB:AB007869; PIDN:AAK86311.1; PID:G15155427; GSPDB:GN00169  
C:Genetics:  
A:Gene: AGR\_C 881  
A:Map position: circular chromosome

Query Match 0.7%; Score 8; DB 2; Length 356;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1155 LLALLVLA 1162  
|||||||  
DB 323 LLALLVLA 330

RESULT 26  
A85016  
hypothetical protein AT4g01220 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: A85016  
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp  
Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488; PMID:10617198  
A:Accession: A85016  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-360 <STO>  
A:Cross-references: GB:NC\_001268; NID:G7267619; PIDN:CAB80931.1; GSPDB:GN00140  
C:Genetics:  
A:Gene: AT4g01220  
A:Map position: 4



Query Match 0.7%; Score 8; DB 2; Length 360;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160  
|||||  
Db 44 LLLALLLV 51

RESULT 27  
S19172  
cytochrome P450 2B4 - rat (fragments)  
N;Alternate names: cytochrome P450 1M2  
N;Contains: oxidoreductase (EC 1.-.-.-)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 22-Nov-1993 #sequence\_revision 21-Jul-1995 #text\_change 04-Mar-2000  
C;Accession: S19172  
R;Yuan, P.M.; Ryan, D.E.; Levin, W.; Shively, J.E.  
Proc. Natl. Acad. Sci. U.S.A. 80, 1169-1173, 1983  
A;Title: Identification and localization of amino acid substitutions between two phenob  
A;Reference number: S19172; MUID:83144040; PMID:6572377  
A;Accession: S19172  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-158;159-200;201-310;311-367 <YUA>  
A;Experimental source: strain Long-Evans  
C;Genetics:  
A;Gene: CYP2B4  
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
C;Keywords: cytochrome; heme; iron; metalloprotein; microsome; monooxygenase; oxidore  
F;312/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 0.7%; Score 8; DB 2; Length 367;  
Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160  
|||||  
Db 6 LLLALLLV 13

RESULT 28  
DH004  
dopamine receptor D4 - human  
C;Species: Homo sapiens (man)  
C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 02-Sep-1997  
C;Accession: S15079  
R;van Tol, H.H.M.; Bunzow, J.R.; Guan, H.C.; Sunahara, R.K.; Seeman, P.; Niznik, H.B.; C  
Nature 350, 610-614, 1991  
A;Title: Cloning of the gene for a human dopamine D(4) receptor with high affinity for d  
A;Reference number: S15079; MUID:91204054; PMID:1840645  
A;Accession: S15079  
A;Molecule type: DNA  
A;Residues: 1-387 <VAN>  
A;Cross-references: EMBL:X58497  
C;Genetics:  
A;Gene: DRD4  
A;Cross-references: GDB:127782; OMIM:126452  
A;Map position: 11p15.5-11p15.5  
A;Introns: 95/3; 133/2; 269/2; 321/1  
C;Superfamily: vertebrate rhodopsin  
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; neurotransmi  
F;34-60/Domain: transmembrane #status predicted <TM1>  
F;72-96/Domain: transmembrane #status predicted <TM2>  
F;110-131/Domain: transmembrane #status predicted <TM3>  
F;153-174/Domain: transmembrane #status predicted <TM4>  
F;192-214/Domain: transmembrane #status predicted <TM5>  
F;215-314/Domain: intracellular #status predicted <INT>  
F;315-339/Domain: transmembrane #status predicted <TM6>  
F;349-368/Domain: transmembrane #status predicted <TM7>  
F;3/Binding site: carbonylate (Asn) (covalent) #status predicted  
F;108-185/Disulfide bonds: #status predicted  
F;149,239/Binding site: phosphate (Ser) (covalent) #status predicted

F;297,306/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 0.7%; Score 8; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLLV 1161  
|||||  
Db 81 LLLALLLV 88

RESULT 29  
S23860  
chloramphenicol resistance protein homolog opdE - Pseudomonas aeruginosa  
C;Species: Pseudomonas aeruginosa  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 31-Dec-2000  
C;Accession: S23860; H83368  
R;Huang, H.; Siebel, R.J.; Bellido, P.; Rawling, E.; Hancock, R.E.W.  
submitted to the EMBL Data Library, July 1992  
A;Description: Analysis of two gene regions involved in the expression of the imipene  
A;Reference number: S23859  
A;Accession: S23860  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-402 <HUR>  
A;Cross-references: EMBL:Z14064; NID:945366; PIDN:CAA78446.1; PID:945368  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; I  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: H83368  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-402 <STO>  
A;Cross-references: GB:AE004648; GB:AE004091; NID:99948237; PIDN:AAG05607.1; GSPDB:GN  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: opdE; PA2219  
C;Superfamily: Streptomyces lividans chloramphenicol resistance protein  
C;Keywords: antibiotic resistance; transmembrane protein

Query Match 0.7%; Score 8; DB 1; Length 402;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1147 GSTLGLL 1154  
|||||  
Db 360 GSTLGLL 367

RESULT 30  
G83568  
probable permease of ABC transporter PA0605 [imported] - Pseudomonas aeruginosa (stra  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C;Accession: G83568  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; I  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: G83568  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-415 <STO>  
A;Cross-references: GB:AE004497; GB:AE004091; NID:99946478; PIDN:AAG03994.1; GSPDB:GN  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: PA0605

Query Match 0.7%; Score 8; DB 2; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLLA 1157  
 |||||  
 Db 388 LGGLLLLA 395

RESULT 31  
 AG3504  
 dihydrofolate synthase (EC 6.3.2.12) [imported] - Brucella melitensis (strain 16M)  
 C:Species: Brucella melitensis  
 C>Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 19-Apr-2002  
 C:Accession: AG3504  
 R:DelVecchio, V.G.; Kapur, V.; Redkar, R.J.; Patra, G.; Majer, C.; Los, T.; Ivanova,  
 ; Mazur, M.; Goltzman, E.; Selkov, E.; Elizek, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens  
 A:Reference number: AD3252; PMID:11756688  
 A:Accession: AG3504  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-442 <KUR>  
 A:Cross-references: GB:AE008917; PIDN:AAL53202.1; PID:gl7984076; GSPDB:GN00190  
 A:Experimental source: strain 16M  
 C:Genetics:  
 A:Gene: BMEI2021  
 A:Map position: 1  
 C:Superfamily: folypolyglutamate synthase  
 C:Keywords: ligase

Query Match 0.7%; Score 8; DB 2; Length 442;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 615 NEDGLIDL 622  
 |||||  
 Db 241 NEDGLIDL 248

RESULT 32  
 T01721  
 hypothetical protein A\_IG002N01.8 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 24-Mar-1999  
 C:Accession: T01721  
 R:Scheet, P.; Maggi, L.  
 Submitted to the EMBL Data Library, June 1997  
 A:Description: The sequence of A. thaliana IG002N01.  
 A:Reference number: Z14407  
 A:Accession: T01721  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-444 <SCH>  
 A:Cross-references: EMBL:AF007269; NID:g2191126; PID:g2191131  
 A:Experimental source: cultivar Columbia  
 C:Genetics:  
 A:Map position: 4  
 A:Introns: 68/1; 235/3; 302/3; 374/3  
 A:Note: A\_IG002N01.8

Query Match 0.7%; Score 8; DB 2; Length 444;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160  
 |||||  
 Db 115 LLLALLLV 122

RESULT 33  
 T47107

benzaldehyde dehydrogenase (NAD) (EC 1.2.1.28) [validated] - Pseudomonas putida plasm  
 C:Species: Pseudomonas putida  
 C>Date: 17-Mar-2000 #sequence\_revision 17-Mar-2000 #text\_change 03-Jun-2002  
 C:Accession: T47107; S13388  
 R:Inoue, J.; Shaw, J.P.; Rexik, M.; Harayama, S.  
 J. Bacteriol. 177, 1196-1201, 1995  
 A:Title: Overlapping substrate specificity of benzaldehyde dehydrogenase (the xylC ge  
 Pseudomonas putida.  
 A:Reference number: Z24352; MUID:95173094; PMID:7868591  
 A:Accession: T47107  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-487 <INO>  
 A:Cross-references: EMBL:U15151; NID:g555994; PIDN:AAA66218.1; PID:g555995  
 R:Chalmers, R.M.; Keen, J.N.; Fawson, C.A.  
 Biochem. J. 273, 99-107, 1991  
 A:Title: Comparison of benzyl alcohol dehydrogenases and benzaldehyde dehydrogenases  
 uene pathway in Pseudomonas putida.  
 A:Reference number: S13398; MUID:9113163; PMID:1989592  
 A:Accession: S13398  
 A:Molecule type: protein  
 A:Residues: 1-44 <CHA>  
 A:Experimental source: strain MT53  
 C:Genetics:  
 A:Gene: xylC  
 A:Genome: plasmid  
 A:Note: TOL plasmid pMW0  
 C:Function:  
 A:Description: EC 1.2.1.28 [validated, MUID:95173094]  
 C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology  
 C:Keywords: oxidoreductase

Query Match 0.7%; Score 8; DB 2; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 626 ALGNVAVIL 633  
 |||||  
 Db 171 ALGNVAVIL 178

RESULT 34  
 O4RTPS  
 cytochrome P450 2B1 - rat  
 N:Alternate names: cytochrome P450 b; cytochrome P450, phenobarbital-inducible  
 N:Contains: unspecific monooxygenase (EC 1.14.14.1)  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 18-Aug-1982 #sequence\_revision 17-May-1996 #text\_change 03-Mar-2000  
 C:Accession: A00176; A54251; A22363; A29298; S03854; A92255; I54796  
 R:Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 2793-2797, 1982  
 A:Title: Primary structure of a cytochrome P-450: coding nucleotide sequence of phenol  
 A:Reference number: A93912; MUID:8222224; PMID:6953431  
 A:Accession: A00176  
 A:Molecule type: mRNA  
 A:Residues: 6-491 <FUJ>  
 A:Cross-references: EMBL:J00719; NID:g203752; PIDN:AAA41024.1; PID:g203753  
 A:Note: the authors translated the codon GAT for residue 166 as Glu, CTC for residue  
 R:Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 5443, 1982  
 A:Title: Primary structure of a cytochrome P450: coding nucleotide sequence of phenob.  
 A:Reference number: A93925  
 A:Contents: annotation  
 A:Note: the mistranslations shown in reference A93912 are acknowledged  
 R:Roberts, E.S.; Hopkins, N.E.; Zaluzec, E.J.; Gage, D.A.; Alworth, W.L.; Hollenberg,  
 Biochemistry 31, 3766-3771, 1994  
 A:Title: Identification of active-site peptides from (3)H-labeled 2-ethynynaphthalen  
 A:Reference number: A54251; MUID:94190899; PMID:8142377  
 A:Accession: A54251  
 A:Molecule type: protein  
 A:Residues: 290-301, 'X' <ROB>  
 R:Suwa, Y.; Mizukami, Y.; Sogawa, K.; Fujii-Kuriyama, Y.  
 J. Biol. Chem. 260, 7980-7984, 1985

A:Title: Gene structure of a major form of phenobarbital-inducible cytochrome P-450 in z  
A:Reference number: A22363; MUID:85234490; PMID:2989270  
A:Accession: A22363  
A:Molecule type: DNA  
A:Residues: 1-91, 'P', 93-204, 'R', 328-356, 'H', 358-391, 'R', 393-415, 'V', 417-433,  
A:Cross-references: GB:100320; NID:G203816; PIDN:AAA41046.1; PID:G203818  
A:Note: the authors translated the codon CAG for residue 57 as Gly, CCG for residue 92 as  
as Glu, AAA for residue 236 as Leu, AGC for residue 259 as Asn, GTT for residue 328 as T  
as Arg  
R:Rangarajan, P.N.; Ravishanker, H.; Padmanaban, G.  
Biochem. Biophys. Res. Commun. 144, 258-263, 1987  
A:Title: Isolation of a cytochrome P-450 gene variant and characterization of its 5' fl  
A:Reference number: A29298; MUID:87213174; PMID:3579906  
A:Accession: A29298  
A>Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-57 <RAN>  
R:Oesch, F.; Waxman, D.J.; Morrissey, J.J.; Honscha, W.; Kissel, W.; Friedberg, T.  
Arch. Biochem. Biophys. 270, 23-32, 1989  
A:Title: Antibodies targeted against hypervariable and constant regions of cytochromes P  
A:Reference number: S03854; MUID:89192373; PMID:2539047  
A:Accession: S03854  
A>Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-18, 146-160, 'E', 162-165, 166, 330-361, 362-380, 402-423 <OES>  
R:Botelho, L.H.; Ryan, D.E.; Levin, W.  
J. Biol. Chem. 254, 5635-5640, 1979  
A:Title: Amino acid compositions and partial amino acid sequences of three highly purified  
or 3-methylcholanthrene.  
A:Reference number: A92255; MUID:79194111; PMID:109438  
A:Accession: A92255  
A:Molecule type: protein  
A:Residues: 1-3, 'T', 5-22 <BOT>  
R:Fujii-Kuriyama, Y.; Mizukami, Y.; Taniguchi, T.; Muramatsu, M.  
Int. Symp. Princess Takamatsu Cancer Res. Fund 12, 31-40, 1982  
A:Title: Molecular cloning and coding nucleotide sequence of complementary DNA of cytoch  
A:Reference number: 154796; MUID:83160754; PMID:6300027  
A:Accession: 154796  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 6-491 <RES>  
A:Cross-references: GB:X37134; NID:G203784; PIDN:AAC42028.1; PID:G203785  
C:Genetics:  
A:Gene: CYP2B1  
A:Inserts: 57/3; 112/1; 162/1; 215/3; 274/3; 322/1; 384/3; 432/1  
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo  
P:295-458/Domain: cytochrome P450 homology <P45>  
P:302/Active site: Thr #status predicted  
P:436/Binding site: heme iron (Cys) (axial ligand) #status predicted  
Query Match 0.74; Score 8; DB 1; Length 491;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1153 LLLALLLV 1160  
Db 6 LLLALLLV 13  
RESULT 35  
O4RT2  
Cytochrome P450 2B2 - rat  
N:Alternate names: cytochrome P450 PB-4; cytochrome P450, phenobarbital-inducible; cyto  
N:Contains: oxidoreductase (EC 1.-.-.-)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 04-Dec-1986 #sequence revision 17-May-1996 #text change 01-Dec-2000  
A:Accession: A21162; A00177; B00176; B92255; S15589; A21872; A32736; S03855; I59060  
R:Mizukami, Y.; Sogawa, K.; Suwa, Y.; Muramatsu, M.; Fujii-Kuriyama, Y.  
Proc. Natl. Acad. Sci. U.S.A. 80, 3958-3962, 1983  
A:Title: Gene structure of a phenobarbital-inducible cytochrome P-450 in rat liver.  
A:Reference number: A21162; MUID:83247397; PMID:6306654  
A:Accession: A21162

A:Molecule type: DNA  
A:Residues: 1-472, 'M', 474-491 <MIZ>  
A:Cross-references: EMBL:J00728; NID:G203845; PIDN:AAA1056.1; PID:G203847  
A:Note: the authors translated the codon AGT for residue 4 as Thr, and ATG for residue  
R:Freij, A.B.; Waxman, D.J.; Kreibich, G.  
J. Biol. Chem. 260, 15253-15265, 1985  
A:Title: The structure of phenobarbital-inducible rat liver cytochrome P-450 isoform  
A:Reference number: A00177; MUID:86059379; PMID:3877725  
A:Accession: A00177  
A:Molecule type: protein  
A:Residues: 1-291, 'P', 293-320, 'AB', 323-475, 'D', 477-491 <FRE>  
R:Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.  
Proc. Natl. Acad. Sci. U.S.A. 79, 2793-2797, 1982  
A:Title: Primary structure of a cytochrome P-450: coding nucleotide sequence of phen  
A:Reference number: A93913; MUID:82222224; PMID:6953431  
A:Accession: B00176  
A:Molecule type: mRNA  
A:Residues: 6-359, 'S', 361-362, 'V', 364-366, 'V', 368-406, 'S', 408-416, 'N', 418, 'A', 420-477  
A:Note: nucleotide sequence for residues 1-5 is not given  
A:Note: the authors translated the codon CAT for residue 166 as Glu, CTG for residue  
R:Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.  
Proc. Natl. Acad. Sci. U.S.A. 79, 5443, 1982  
A:Title: Primary structure of a cytochrome P450: coding nucleotide sequence of phenot  
A:Reference number: A93925  
A:Contents: annotation; revisions  
A:Note: the mistranslations in reference A93912 are acknowledged  
R:Botelho, L.H.; Ryan, D.E.; Levin, W.  
J. Biol. Chem. 254, 5635-5640, 1979  
A:Title: Amino acid compositions and partial amino acid sequences of three highly pur  
or 3-methylcholanthrene.  
A:Reference number: A92255; MUID:79194111; PMID:109438  
A:Accession: B92255  
A:Molecule type: protein  
A:Residues: 1-3, 'T', 5-22 <BOT>  
R:Lacroix, D.; Desrochers, M.; Lambert, M.; Anderson, A.  
Gene 86, 201-207, 1990  
A:Title: Alternative splicing of mRNA encoding rat liver cytochrome P450e (P450IIB2).  
A:Reference number: S15589; MUID:90215299; PMID:2323573  
A:Accession: S15589  
A:Molecule type: mRNA  
A:Residues: 105-113, 'F', 115-274, 'VSPAMRE', 275-321, 'E', 323-491 <LAC>  
A:Cross-references: EMBL:M34452; NID:G203679; PIDN:AAA41004.1; PID:G203680  
A:Note: translation of the nucleotide sequence is not complete  
R:Philips, I.R.; Shephard, E.A.; Ashworth, A.; Rabin, B.R.  
Gene 24, 41-52, 1983  
A:Reference number: A21872  
A:Accession: A21872  
A:Molecule type: mRNA  
A:Residues: 168-321, 'E', 323-443, 'K', 445-491 <PHI>  
R:Affolter, M.; Anderson, A.  
Biochem. Biophys. Res. Commun. 118, 655-662, 1984  
A:Title: Segmental homologies in the coding and 3' non-coding sequences of rat liver  
A:Reference number: A32736; MUID:84153837; PMID:6322758  
A:Accession: A32736  
A:Molecule type: mRNA  
A:Residues: 385-491 <APP>  
A:Cross-references: GB:K01626; NID:G203782; PIDN:AAA41037.1; PID:G203783  
R:Oesch, F.; Waxman, D.J.; Morrissey, J.J.; Honscha, W.; Kissel, W.; Friedberg, T.  
Arch. Biochem. Biophys. 270, 23-32, 1989  
A:Title: Antibodies targeted against hypervariable and constant regions of cytochrome  
A:Reference number: S03854; MUID:89192373; PMID:2539047  
A:Accession: S03855  
A>Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 329-358, 'AS', 361, 362, 363-380, 402-423 <OES>  
R:Atchison, M.L.; Adenik, M.  
Proc. Natl. Acad. Sci. U.S.A. 83, 2300-2304, 1986  
A:Title: Gene conversion in a cytochrome P-450 gene family.  
A:Reference number: I59060; MUID:86205943; PMID:3458196  
A:Accession: I59060  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 323-431 <RES>

A;Cross-references: GB:M13234; NID:g203848; PIDN:AAA41057.1; PID:g554434

C;Genetics:

A;Introns: 384/3

C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology

C;Keywords: alternative splicing; chromoprotein; electron transfer; endoplasmic reticulum

F:295-458/Domain: cytochrome P450 homology <P45>

F:436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 0.7%; Score 8; DB 1; Length 491;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 LLLLLLLV 1160

Db 6 LLLLLLLV 13

RESULT 36

B31047

testosterone 16alpha-hydroxylase (EC 1.14.14.-) cytochrome P450 2B10 - mouse

N;Alternate names: cytochrome P450CDB; cytochrome P450pf3/46

C;Species: Mus musculus (house mouse)

C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 04-Mar-2000

C;Accession: B31047; A60559

R;Noshiro, M.; Lakso, M.; Kawajiri, K.; Negishi, M.

Biochemistry 27, 6434-6443, 1988

A;Title: Rip locus: regulation of female-specific isozyme (I-P-450-16-alpha) of testosterone

A;Reference number: A31047; MUID:89118235; PMID:3219345

A;Accession: B31047

A;Molecule type: mRNA

A;Residues: 1-500 <NO2>

A;Cross-references: EMBL:M21856; NID:g201968; PIDN:AAA40425.1; PID:g201969

A;Note: the authors translated the codon TTA for residue 55 as Phe, AGA for residue 133

as Val, AAA for residue 236 as Asn, AGG for residue 251 as Lys, AAG for residue 253 as

3 as Asp

A;Note: the authors translated the codon GCA for residue 281 as Thr, TTC for residue 283

8 as Arg, ACA for residue 321 as Ala, GTG for residue 331 as Leu, TCA for residue 349 as

A;Note: the sequence nucleotide translation from Fig. 8 is inconsistent with the nucleot

R;Bornheim, L.M.; Correia, M.A.

Mol. Pharmacol. 36, 377-383, 1989

A;Title: Purification and characterization of a mouse liver cytochrome P-450 induced by

A;Reference number: A60559; MUID:89384475; PMID:2779523

A;Accession: A60559

A;Molecule type: protein

A;Residues: 1-15 <BOR>

A;Note: this enzyme was induced by cannabidiol

C;Genetics:

A;Gene: Cyp2b-10

C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology

C;Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo

F:295-467/Domain: cytochrome P450 homology <P45>

F:445/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 0.7%; Score 8; DB 2; Length 500;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 LLLLLLLV 1160

Db 6 LLLLLLLV 13

RESULT 37

AC1314

malolactic enzyme (malate dehydrogenase) homolog lmo1915 [imported] - Listeria monocytog

C;Species: Listeria monocytogenes

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001

C;Accession: AC1314

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;

ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AC1314

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-547 <GLA>

A;Cross-references: GB:NC\_003210; PIDN:CAC99993.1; PID:g16411368; GSPDB:GN00177

A;Experimental source: strain 5GD-e

C;Genetics:

A;Gene: lmo1915

C;Superfamily: malate dehydrogenase (oxaloacetate-decarboxylating)

Query Match 0.7%; Score 8; DB 2; Length 547;

Best Local Similarity 100.0%; Pred. No. 39;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 266 MIVITDGE 273

Db 141 MIVITDGE 148

RESULT 38

AC1686

malolactic enzyme (malate dehydrogenase) homolog lin2029 [imported] - Listeria innocu

C;Species: Listeria innocua

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001

C;Accession: AC1686

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;

ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AC1686

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-547 <GLA>

A;Cross-references: GB:AL592022; PIDN:CAC97359.1; PID:g16414530; GSPDB:GN00178

A;Experimental source: strain Clp11262

C;Genetics:

A;Gene: lin2029

C;Superfamily: malate dehydrogenase (oxaloacetate-decarboxylating)

Query Match 0.7%; Score 8; DB 2; Length 547;

Best Local Similarity 100.0%; Pred. No. 39;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 266 MIVITDGE 273

Db 141 MIVITDGE 148

RESULT 39

NIWEEB

E1 protein - bovine papillomavirus type 1

C;Species: bovine papillomavirus type 1

C;Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 28-Jul-2000

C;Accession: A03663

R;Chen, E.Y.; Howley, P.M.; Levinson, A.D.; Seeburg, P.H.

Nature 299, 529-534, 1982

A;Title: The primary structure and genetic organization of the bovine papillomavirus

A;Reference number: A93289; MUID:81012974; PMID:6289124

A;Accession: A03663

A;Molecule type: DNA

A;Residues: 1-605 <CHE>

A;Cross-references: GB:X02346; GB:J02044; GB:M24622; GB:X00473; NID:g60965; PIDN:CAB4

C;Superfamily: papillomavirus E1 protein

C;Keywords: early protein

Query Match 0.7%; Score 8; DB 1; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 927 LAAGSDSN 934  
 |||||  
 Db 337 LAAGSDSN 344

RESULT 40  
 A72429  
 oligopeptide ABC transporter, periplasmic oligopeptide-binding protein - Thermotoga maritima  
 C:Species: Thermotoga maritima  
 C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
 C:Accession: A72429  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.N.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.  
 Nature 399, 323-329, 1999  
 A>Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing  
 A:Reference number: A72200; PMID:10360571  
 A:Accession: A72429  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-606 <ARN>  
 A:Cross-references: GB:AE001690; GS:AE000512; NID:g4980496; PIDN:AAD35125.1; PID:g498051  
 A:Experimental source: strain MSB8  
 C:Genetics:  
 A:Gene: TM0031  
 C:Superfamily: dipeptide transport protein

Query Match 0.7%; Score 8; DB 2; Length 606;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLV 1161  
 |||||  
 Db 9 LLLALLV 16

RESULT 41  
 A69845  
 Na+/H+ antiporter homolog yjbQ - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999  
 C:Accession: A69845  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertoni, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chao, A.; Ehrlich, S.D.; Emerson, P.T.; Estian, K.D.; Evington, J.; Fabbet, C.; Ferrari, E.; Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, Rech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; PMID:98044033; PMID:9384377  
 A:Accession: A69845  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-614 <KUN>  
 A:Cross-references: GB:Z99110; GB:AL009126; NID:g2633472; PIDN:CAB13021.1; PID:el183184;  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yjbQ

Query Match 0.7%; Score 8; DB 2; Length 614;

Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160  
 |||||  
 Db 313 LLLALLV 320

RESULT 42  
 W1WLB2  
 E1 protein - bovine papillomavirus type 2  
 C:Species: bovine papillomavirus type 2  
 C>Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 12-Jun-1998  
 C:Accession: C31169  
 R:Groff, D.E.; Mitra, R.; Lancaster, W.D.  
 Submitted to GenBank, May 1988  
 A:Reference number: A94519  
 A:Accession: C31169  
 A:Molecule type: DNA  
 A:Residues: 1-620 <GRO>  
 A:Cross-references: GB:M20219; GS:M19551; NID:g332996  
 C:Superfamily: papillomavirus E1 protein  
 C:Keywords: early protein; glycoprotein  
 F:72,109,173/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 0.7%; Score 8; DB 1; Length 620;  
 Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 927 LAAGSDSN 934  
 |||||  
 Db 352 LAAGSDSN 359

RESULT 43  
 AC0527  
 ferrichrome transport protein FhuB precursor STV0221 (imported) - Salmonella enterica  
 C:Species: salmonella enterica subsp. enterica serovar Typhi  
 A>Note: this species has also been called Salmonella typhi  
 C>Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C:Accession: AC0527  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Park, S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, A.; Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi  
 A:Reference number: AB0502; PMID:21534947; PMID:11677608  
 A:Accession: AC0527  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-685 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD01355.1; PID:g16501482; GSPDB:GN00176  
 C:Genetics:  
 A:Gene: STV0221  
 C:Superfamily: vitamin B12 transport protein btuC

Query Match 0.7%; Score 8; DB 2; Length 685;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160  
 |||||  
 Db 36 LLLALLV 43

RESULT 44  
 T16306  
 hypothetical protein F40F4.7 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T16306  
 R:Wilson, R.

submitted to the EMBL Data Library, November 1995  
 A:Description: The sequence of *C. elegans* cosmid F40F4.  
 A:Reference number: Z18493

A:Accession: T16306  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-697 <WIL>  
 A:Cross-references: EMBL:U40420; NID:G1065513; PID:G1065515; PIDN:AAA81431.1; CESP:F40F4  
 C:Gene: CESP:F40F4.7  
 A:Introns: 14/2; 86/3; 135/3; 223/1; 244/1; 331/1; 365/1; 396/1; 435/1; 559/3; 627/2; 65

Query Match 0.7%; Score 8; DB 2; Length 697;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 557 YNDVVVGA 564  
 DB 593 YNDVVVGA 600

## RESULT 45

A88445

protein C26E6.10 [imported] - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001

A:Accession: A88445  
 R:Anonymous, The *C. elegans* Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology  
 A:Reference number: A75000; PMID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: A88445

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-739 <STO>

A:Cross-references: GB:chr\_III; PIDN:AAA21163.1; PID:G532810; GSPDB:GN00021; CESP:C26E6.

A:Gene: C26E6.10

C:Genetics:

A:Map position: 3

Query Match 0.7%; Score 8; DB 2; Length 739;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 430 TSVVSSRQ 437  
 DB 234 TSVVSSRQ 241

## RESULT 46

E83349

probable ClpA/B-type proteinase PA2371 [imported] - *Pseudomonas aeruginosa* (strain PA01)  
 C:Species: *Pseudomonas aeruginosa*

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: E83349

R:Stover, C.X.; Phan, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim

; Lory, S.; Olson, M.V.

Nature 405, 959-964, 2000

A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
 A:Reference number: A82950; PMID:20437337; PMID:10984043

A:Accession: E83349

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-849 <STO>

A:Cross-references: GB:AE004663; GB:AE004091; NID:G9948405; PIDN:AG05759.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA2371

C:Superfamily: endopeptidase Clp ATP-binding chain

Query Match 0.7%; Score 8; DB 2; Length 849;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159  
 DB 111 GLLLLALL 118

## RESULT 47

T31437

integrin alpha chain SU2 - sea urchin (*Lytechinus variegatus*)  
 C:Species: *Lytechinus variegatus* (variegated urchin)

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jan-2000

C:Accession: T31437

R:Hertzler, P.L.; McClay, D.R.

submitted to the EMBL Data Library, May 1998

A:Description: Alpha SU2, a sea urchin integrin which binds laminin.

A:Reference number: Z21035

A:Accession: T31437

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1041 <HER>

A:Cross-references: EMBL:AF067658; NID:G3220240; PID:G3220241; PIDN:AAC23572.1

A:Experimental source: developmental stage embryo

C:Function:

A:Description: binds laminin

C:Superfamily: integrin alpha-2b chain

Query Match 0.7%; Score 8; DB 2; Length 1041;  
 Best Local Similarity 100.0%; Pred. No. 68;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 498 LLVGAPMY 505  
 DB 308 LLVGAPMY 315

## RESULT 48

T25068

hypothetical protein T21C9.2 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T25068

R:McMurray, A.

submitted to the EMBL Data Library, May 1996

A:Reference number: Z19977

A:Accession: T25068

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1065 <WIL>

A:Cross-references: EMBL:Z73098; PIDN:CAA97331.1; GSPDB:GN00023; CESP:T21C9.2

A:Experimental source: Clone T21C9

C:Genetics:

A:Gene: CESP:T21C9.2

A:Map position: 5

A:Introns: 67/2; 147/2; 207/3; 242/3; 293/3; 825/2; 897/1; 1032/2

Query Match 0.7%; Score 8; DB 2; Length 1065;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 866 FASLIQKE 873  
 DB 457 FASLIQKE 464

## RESULT 49

S40311

integrin - fruit fly (*Drosophila melanogaster*)  
 C:Species: *Drosophila melanogaster*

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 21-Jul-2000

C:Accession: S40311

R;Wehrli, M.; DiAntonio, A.; Fearnley, I.M.; Smith, R.J.; Wilcox, M.  
 Mech. Dev. 43, 21-36, 1993  
 A>Title: Cloning and characterization of alpha(P51), a novel Drosophila melanogaster integrin  
 A;Reference number: S40311; MUID:94059764; PMID:8240969  
 A;Accession: S40311  
 A>Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-1146 <WEH>  
 A;Cross-references: EMBL:X73975; NID:g440143; PIDN:CAA52155.1; PID:g440144  
 A;Gene: FlyBase:new  
 C;Genetics:  
 A;Cross-references: FlyBase:FBgn0004456  
 C;Superfamily: Integrin alpha-2b chain

Query Match 0.7%; Score 8; DB 2; Length 1146;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 8; Conservative 0; Mismatches 0; Gaps 0;  
 Indels 0;

Qy 1151 GGLLLALL 1158  
 |||||  
 Db 1094 GGLLLALL 1101

RESULT 50  
 RWHUIC  
 cell surface glycoprotein CD11b precursor [validated] - human  
 N;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Mac-1  
 eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain  
 C;Species: Homo sapiens (man)  
 C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000  
 C;Accession: A31108; A26915; R41600; A30892; A32218; A46526; A36091; I52567  
 R;Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.  
 J. Cell Biol. 106, 2153-2158, 1988  
 A;Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b)  
 A;Reference number: A31108; MUID:88315033; PMID:2457584  
 A;Accession: A31108  
 A;Molecule type: mRNA  
 A;Residues: 1-1153 <COR>  
 A;Cross-references: GB:J03925; NID:g187284; PIDN:AA59544.1; PID:g307148  
 A;Note: part of this sequence was confirmed by protein sequencing  
 R;Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.  
 J. Cell Biol. 106, 2153-2158, 1988  
 A;Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor M  
 A;Reference number: A28915; MUID:88257215; PMID:2454931  
 A;Accession: A28915  
 A;Molecule type: mRNA  
 A;Residues: 1-499,501-965, 'P', 967-1153 <ARN>  
 A;Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:g186935; PIDN:AA594  
 A;Note: the authors translated the codon TAC for residue 1129 as Thr  
 A;Note: part of this sequence, including the amino end of the mature protein, was confirmed by protein sequencing  
 R;Shelley, C.S.; Arnaout, M.A.  
 Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991  
 A;Title: The promoter of the CD11b gene directs myeloid-specific and developmentally regulated expression  
 A;Reference number: A41600; MUID:92073318; PMID:1683702  
 A;Accession: A41600  
 A;Molecule type: DNA  
 A;Residues: 1-9 <SHE>  
 A;Cross-references: GB:M76724; NID:g180018; PIDN:AA58410.1; PID:g553215  
 R;Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988  
 A;Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesion receptor  
 A;Reference number: A94193; MUID:88190151; PMID:2833753  
 A;Accession: A30892  
 A;Molecule type: mRNA  
 A;Residues: 917-1042 <AR2>  
 A;Cross-references: GB:M18044  
 R;Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989  
 A;Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor  
 A;Reference number: A32218; MUID:89098893; PMID:2563162  
 A;Accession: A32218  
 A;Molecule type: mRNA

A;Residues: 9-1153 <HIC>  
 A;Cross-references: GB:J04145; NID:g189068; PIDN:AA559903.1; PID:g386975  
 A;Note: part of this sequence was confirmed by protein sequencing  
 R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.  
 J. Immunol. 150, 480-490, 1993  
 A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alphaM chain during evolution.  
 A;Reference number: A46526; MUID:93123748; PMID:8419480  
 A;Accession: A46526  
 A>Status: not compared with conceptual translation  
 A;Molecule type: DNA  
 A;Residues: 1-499,501-1153 <FLB>  
 A;Cross-references: GB:SS2227; NID:g263047; PIDN:AA524821.1; PID:g263049  
 A;Note: the last three bases of intron 13, CAG, are included in some but not all matings  
 R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.P.; Arnaout, M.A.  
 Biochim. Biophys. Acta 874, 368-371, 1986  
 A;Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across species  
 A;Reference number: A90664; MUID:87076671; PMID:3539202  
 A;Accession: A26091  
 A;Molecule type: protein  
 A;Residues: 17-31 <PIB>  
 A;Experimental source: granulocytes  
 R;Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.  
 Blood 79, 865-870, 1992  
 A;Title: Characterization of the myeloid-specific CD11b promoter.  
 A;Reference number: I52567; MUID:92144986; PMID:1346576  
 A;Accession: I52567  
 A>Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-9 <RES>  
 A;Cross-references: GB:M84477; NID:g180184; PIDN:AA51960.1; PID:g553219  
 C;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1  
 C;Genetics:  
 A;Gene: GDB:ITGAM; CR3A  
 A;Cross-references: GDB:120599; OMIM:120980  
 A;Map position: 16p11.2-16p11.2  
 A;Note: promoter contains a GATA motif and two Spi consensus binding sites  
 C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat 1  
 C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer;  
 F17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>  
 F17-1108/Domain: extracellular #status predicted <EXT>  
 F148-318/Domain: von Willebrand factor type A repeat homology <VWA2>  
 F1465-473/Region: calcium/magnesium binding #status predicted  
 F530-538/Region: calcium/magnesium binding #status predicted  
 F593-601/Region: calcium/magnesium binding #status predicted  
 F1109-1134/Domain: transmembrane #status predicted <TM>  
 F1195-1153/Domain: intracellular #status predicted <INT>  
 P:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Bind

Query Match 0.7%; Score 8; DB 1; Length 1153;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1151 GGLLLALL 1158  
 |||||  
 Db 1116 GGLLLALL 1123

RESULT 51  
 RWHUIC  
 cell surface glycoprotein CD11c precursor - human  
 N;Alternate names: leukocyte adhesion receptor p150,95 alpha chain  
 C;Species: Homo sapiens (man)  
 C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 22-Jun-1999  
 C;Accession: A36584; A35543; S00864  
 R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.  
 J. Biol. Chem. 265, 12750-12751, 1990  
 A;Reference number: A36584  
 A;Contents: erratum  
 A;Accession: A36584  
 A;Molecule type: DNA



A;Residues: 1-1163 <COR>  
A;Note: this revision to the sequence from reference A35543 includes the carboxyl end  
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.  
J. Biol. Chem. 265, 2782-2786, 1990  
A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.  
A;Reference number: A35543; MUID:90153906; PMID:2303426  
A;Accession: A35543  
A;Molecule type: DNA  
A;Residues: 1-834 <CO2>  
A;Note: this sequence has been revised in reference A3584  
R;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.  
EMBO J. 6, 4023-4028, 1987  
A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte  
integrin alpha subunit, the leukocyte p150,95 molecule.  
A;Reference number: S00864; MUID:88166645; PMID:3327687  
A;Accession: S00864  
A;Molecule type: mRNA  
A;Residues: 1-755, 'L', 757-1163 <CO3>  
A;Cross-references: GB:M81695; EMBL:Y00093; NID:9487829; PIDN:AAAS9180.1; PID:9487830  
A;Note: part of this sequence was confirmed by protein sequencing  
A;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my  
C;Genetics:  
A;Gene: GDB:ITGAX; CD11C  
A;Cross-references: GDB:119758; OMIM:151510  
A;Map position: 16p11.2-16p11.2  
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homol  
C;Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;1-19/Product: cell surface glycoprotein CD11c #status predicted <MAT>  
F;20-1107/Domain: extracellular #status predicted <EXT>  
F;149-319/Domain: von Willebrand factor type A repeat homology <VWA4>  
F;1108-1133/Domain: transmembrane #status predicted <TM>  
F;1134-1163/Domain: intracellular #status predicted <INT>  
F;61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre  
Query Match 0.7%; Score 8; DB 1; Length 1163;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1151 GGLLLAL 1158  
Db 1115 GGLLLAL 1122  
RESULT 52  
A53213  
Integrin alpha-B chain - human  
C;Species: Homo sapiens (man)  
C;Date: 19-Oct-1995 #sequence\_revision 31-Mar-2001 #text\_change 31-Mar-2001  
C;Accession: A53213  
R;Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.  
J. Biol. Chem. 269, 6015-6025, 1994  
A;Title: Molecular cloning of the human mucosal lymphocyte integrin alpha(B) subunit. Un  
A;Reference number: A53213; MUID:94164962; PMID:8119947  
A;Accession: A53213  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1179 <SHA>  
A;Cross-references: GB:L25851; NID:9457244; PID:9457245  
C;Genetics:  
A;Gene: GDB:ITGAE  
A;Cross-references: GDB:330801  
A;Map position: 17p13  
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homol  
F;1199-371/Domain: von Willebrand factor type A repeat homology <VWA3>  
Query Match 0.7%; Score 8; DB 2; Length 1179;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 380 VLLGAVGA 387  
Db 416 VLLGAVGA 423  
RESULT 53  
S57664  
IGA-specific metalloendopeptidase homolog sepA precursor - Shigella flexneri  
C;Species: Shigella flexneri  
C;Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 08-Dec-2000  
C;Accession: S57664; S69769; S69768  
R;Benjelloun-Touimi, Z.; Sansonetti, P.J.; Parsot, C.  
submitted to the EMBL Data Library, February 1995  
A;Description: Characterization of SepA, the major extracellular protein of Shigella  
A;Reference number: S57664  
A;Accession: S57664  
A;Molecule type: DNA  
A;Residues: 1-1366 <BEN>  
A;Cross-references: EMBL:Z48219; NID:9886952; PIDN:CAA88252.1; PID:9886953  
R;Benjelloun-Touimi, Z.; Sansonetti, P.J.; Parsot, C.  
Mol. Microbiol. 17, 123-135, 1995  
A;Title: SepA, the major extracellular protein of Shigella flexneri: autonomous secre  
A;Reference number: S69768; MUID:96020667; PMID:7476198  
A;Accession: S69769  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-666,1077-1366 <BEN>  
A;Cross-references: EMBL:Z48219; NID:9886952; PIDN:CAA88252.1; PID:9886953  
A;Experimental source: strain M90F (serotype 5)  
A;Accession: S69768  
A;Molecule type: protein  
A;Residues: 57-61, 'Q', 63-72, 'X', 544-557, 'X', 1057-1068 <BEP>  
A;Note: 6-Glu, 14-Thr, 18-Gly and 26-Gln were also found  
C;Genetics:  
A;Gene: sepA  
C;Superfamily: IGA-specific metalloendopeptidase  
C;Keywords: extracellular protein  
F;1-56/Domain: signal sequence #status predicted <SIG>  
F;57-1366/Product: IGA-specific metalloendopeptidase homolog sepA #status predicted <  
Query Match 0.7%; Score 8; DB 2; Length 1366;  
Best Local Similarity 100.0%; Pred. No. 86;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 100 KDNRLGL 107  
Db 1335 KDNRLGL 1342  
RESULT 54  
D81675  
polymorphic membrane protein B/C family TC0695 [imported] - Chlamydia muridarum (stra  
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn  
C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
C;Accession: D81675  
R;Read, T.B.; Bruhan, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.P.; White, O.; Hicke  
C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe  
Nucleic Acids Res. 28, 1397-1406, 2000  
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39  
A;Reference number: A81500; MUID:20150255; PMID:10684935  
A;Accession: D81675  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1460 <TET>  
A;Cross-references: GB:AE002338; GB:AE002160; NID:97190724; PIDN:AAF39511.1; PID:9719  
A;Experimental source: strain Nigg (MoPn)  
C;Genetics:  
A;Gene: TC0695  
Query Match 0.7%; Score 8; DB 2; Length 1460;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 583 SILKTPKQ 590  
Db 1311 SILKTPKQ 1318

RESULT 55  
 S67250  
 DNA-directed RNA polymerase (EC 2.7.7.6) I 190K chain - yeast (*Saccharomyces cerevisiae*)  
 N:Alternate names: DNA-directed RNA polymerase A 190K chain; protein O6276; protein YOR34  
 C:Species: *Saccharomyces cerevisiae*  
 C>Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 20-Jun-2000  
 C:Accession: S67250; A29926; S67396  
 R:Goffeau, A.; Purnelle, B.  
 submitted to the Protein Sequence Database, July 1996  
 A:Reference number: S67246  
 A:Accession: S67250  
 A:Molecule type: DNA  
 A:Residues: 1-1664 <GCF>  
 A:Cross-references: EMBL:Z75249; NID:G1420740; PIDN:CAA98665.1; PID:G1420741; MIPS:YOR34  
 R:Experimental source: strain S288C  
 R:Memet, S.; Gouy, M.; March, C.; Sentenac, A.; Buhler, J.M.  
 J. Biol. Chem. 263, 2830-2839, 1988  
 A:Title: RPA190, the gene coding for the largest subunit of yeast RNA polymerase A.  
 A:Reference number: A29926; MUID:88139333; PMID:2830265  
 A:Accession: A29926  
 A:Molecule type: DNA  
 A:Residues: 1-157, 159-1664 <MEM>  
 A:Cross-references: EMBL:J03530; NID:G172214; PIDN:AAA34890.1; PID:G172215  
 R:Purnelle, B.; Goffeau, A.  
 submitted to the EMBL Data Library, February 1996  
 A:Description: Nucleotide sequence analysis of a 40 kb segment on the right arm of yeast  
 chromosome I genes.  
 A:Reference number: S67392  
 A:Accession: S67396  
 A:Molecule type: DNA  
 A:Residues: 1-1664 <PUR>  
 A:Cross-references: EMBL:X95720; NID:G1199839; PIDN:CAA65029.1; PID:G1199844  
 C:Genetics:  
 A:Gene: SGD:RPA190  
 A:Cross-references: SGD:S0005868; MIPS:YOR341w  
 A:Map position: 15R  
 C:Superfamily: Trypanosoma DNA-directed RNA polymerase I largest chain  
 C:Keywords: DNA binding; nucleotidyltransferase; transcription; zinc finger  
 F:62-78/Region: zinc finger CCH motif  
  
 Query Match 0.7%; Score 8; DB 2; Length 1664;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 481 GSRITSVD 488  
 |||||  
 DB 3 GSRITSVD 15  
  
 RESULT 56  
 H88444  
 protein C26E6.12 [imported] - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
 C:Accession: H88444  
 R:anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology  
 A:Reference number: A75000; MUID:99069613; PMID:9851916  
 A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C\_elegans/  
 A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: H88444  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1802 <STO>  
 A:Cross-references: GB:chr\_III; PIDN:AAA21166.1; PID:G532813; GSPDB:GN00021; CESP:C26E6.12  
 C:Genetics:  
 A:Gene: C26E6.12  
 A:Map position: 3  
  
 Query Match 0.7%; Score 8; DB 2; Length 1802;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 430 TSVVSRQ 437  
 |||||  
 DB 1297 TSVVSRQ 1304  
  
 RESULT 57  
 RRVUBY  
 genome polyprotein - bunyamwera virus  
 N:Alternate names: L protein  
 N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
 C:Species: bunyamwera virus  
 C>Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 11-Jun-1999  
 C:Accession: A33744  
 R:Elliot, R.M.  
 Virology 173, 426-436, 1989  
 A:Title: Nucleotide sequence analysis of the large (L) genomic RNA segment of Bunyamwera virus  
 A:Reference number: A33744; MUID:90085791; PMID:5596023  
 A:Accession: A33744  
 A:Molecule type: Genomic RNA  
 A:Residues: 1-2238 <ELL>  
 A:Cross-references: GB:X14383; NID:G58712; PIDN:CAA32553.1; PID:G58713  
 C:Genetics:  
 A:Map position: segment L  
 C:Superfamily: bunyamwera RNA-directed RNA polymerase  
 C:Keywords: nucleotidyltransferase  
  
 Query Match 0.7%; Score 8; DB 1; Length 2238;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 184 FLINILKK 191  
 |||||  
 DB 1333 FLINILKK 1340  
  
 RESULT 58  
 D81666  
 hypothetical protein TC0766 [imported] - *Chlamydia muridarum* (strain Nigg)  
 C:Species: *Chlamydia muridarum*, *Chlamydia trachomatis* MoPn  
 C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
 C:Accession: D81666  
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke, C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; McClarty, J.; McClarty, G.; Salzberg, E.S.; et al.  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A:Title: Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39  
 A:Reference number: A81500; MUID:20150255; PMID:10684935  
 A:Accession: D81666  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-47 <TET>  
 A:Cross-references: GB:AE002345; GB:AE002160; NID:G7190791; PIDN:AAF39569.1; PID:G7190791  
 A:Experimental source: strain Nigg (MoPn)  
 C:Genetics:  
 A:Gene: TC0766  
  
 Query Match 0.6%; Score 7; DB 2; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1153 LLLALL 1159  
 |||||  
 DB 4 LLLALL 10  
  
 RESULT 59  
 T47376  
 hypothetical protein T5C2.40 - *Arabidopsis thaliana*  
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C:Accession: T47376  
 R:Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Lemcke, K.

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24463  
 A:Accession: T47376  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-74 <OBE>  
 A:Cross-references: EMBL:AL138664  
 A:Experimental source: cultivar Columbia; BAC clone T5C2  
 C:Genetics:  
 A:Map position: 3  
 A:Introns: 36/3; 55/3  
 A>Note: T5C2.40

Query Match 0.6%; Score 7; DB 2; Length 74;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 786 HCVPLV 792  
 Db 13 HCVPLV 19

# RESULT 60

I49515  
 B144 protein B - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
 A:Accession: I49515  
 R:Tsuje, I.; Shen, F.  
 Immunogenetics 26, 378-380, 1987  
 A:Title: A gene in the H-2S: H-2D interval of the major histocompatibility complex which  
 A:Reference number: I49514; MUID:89031493; PMID:3117682  
 A:Accession: I49515  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-89 <RS>  
 A:Cross-references: GB:M18187; NID:g192097; PIDN:AAA37273.1; PID:g192099

Query Match 0.6%; Score 7; DB 2; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 81;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1150 LGGLLLL 1156  
 Db 36 LGGLLLL 42

# RESULT 61

A99774  
 probable excisionase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05099)  
 C:Species: Escherichia coli  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
 A:Accession: A99774  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene  
 A:Reference number: A99629; MUID:21156231; PMID:11258796  
 A:Accession: A99774  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-94 <HAY>  
 A:Cross-references: GB:BA000007; PIDN:BA34584.1; PID:g13360621; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain RIMD 0509952  
 C:Genetics:  
 A:Gene: EC81161

Query Match 0.6%; Score 7; DB 2; Length 94;  
 Best Local Similarity 100.0%; Pred. No. 85;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 637 PVQINA 643  
 Db 15 PVQINA 21

Db 18 PVQINA 24

# RESULT 62

F85636  
 probable excisionase for bacteriophage BP-933W xisW [imported] - Escherichia coli (st  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 A:Accession: F85636  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: F85636  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-99 <STO>  
 A:Cross-references: GB:AB005174; NID:g12514277; PIDN:AAG55554.1; GSPDB:GN00145; UMG:  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: xisW

Query Match 0.6%; Score 7; DB 2; Length 99;  
 Best Local Similarity 100.0%; Pred. No. 89;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 637 PVQINA 643  
 Db 23 PVQINA 29

# RESULT 63

G75257  
 hypothetical protein - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
 A:Accession: G75257  
 R:White, O.; Sisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.  
 M.; Shen, M.O.; Vamathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: G75257  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-103 <WHI>  
 A:Cross-references: GB:AE000513; NID:g6460395; PIDN:AAF12117.1; PID:g6461  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR2573  
 A:Map position: 1

Query Match 0.6%; Score 7; DB 2; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 92;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1157 ALLVIAL 1163  
 Db 15 ALLVIAL 21

# RESULT 64

PH0141  
 integrin alpha chain - chicken (fragments)  
 C:Species: Gallus gallus (chicken)  
 C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 15-Sep-2003  
 A:Accession: PH0141  
 R:Syfrig, J.; Mann, K.; Paulsson, M.  
 Exp. Cell Res. 194, 165-173, 1991  
 A:Title: An abundant chick gizzard integrin is the avian alpha1-beta1 integrin hetero  
 A:Reference number: PH0141; MUID:91224161; PMID:1851093

A:Accession: PH0141  
 A:Molecule type: protein  
 A:Residues: 1-104 <SYF>  
 A:Experimental source: gizzard  
 C:Comment: Integrins are a family of cellular adhesion receptors that mediate interaction  
 C:Keywords: cell adhesion

Query Match 0.6%; Score 7; DB 2; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 93;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 QIGSYFG 481  
 |||||  
 Db 45 QIGSYFG 51

## RESULT 65

S57386

hypothetical protein orf 00954 - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae  
 C:Date: 28-Oct-1995 #sequence\_revision 05-Sep-1996 #text\_change 20-Oct-2000  
 C:Accession: S57386  
 R:Zumstein, E.; Pearson, B.M.; Kalogeropoulos, A.; Schweizer, M.  
 Yeast 11, 975-986, 1995  
 A:Title: A 29.425 kb segment on the left arm of yeast chromosome XV contains more than  
 A:Reference number: S57374; MUID:96021609; PMID:8533473  
 A:Accession: S57386  
 A:Status: nucleic acid sequence not shown; conceptual translation of pseudogene  
 A:Molecule type: DNA  
 A:Residues: 1-106 <ZUM>  
 A:Cross-references: EMBL:X83121; NID:9600461; PIDN:CAA58196.1; PID:9600475  
 C:Comment: There is no evidence that this sequence is expressed.  
 C:Genetics:  
 A:Map position: 15L  
 C:Keywords: pseudogene

Query Match 0.6%; Score 7; DB 4; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 94;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 SFLACSP 123  
 |||||  
 Db 17 SFLACSP 23

## RESULT 66

JC1127

major allergen chain 2 precursor, short form - cat

C:Species: Felis silvestris catus (domestic cat)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 05-Nov-1999  
 C:Accession: JC1127  
 R:Griffith, I.J.; Craig, S.; Pollock, J.; Yu, X.B.; Morgenstern, J.P.; Rogers, B.L.  
 Gene 113, 263-268, 1992  
 A:Title: Expression and genomic structure of the genes encoding Fd1, the major allergen  
 A:Reference number: JC1126; MUID:92241678; PMID:1572548  
 A:Accession: JC1127  
 A:Molecule type: DNA  
 A:Residues: 1-107 <GRI>  
 A:Cross-references: GB:X62478; NID:9395406; PIDN:CAA44345.1; PID:9395407  
 A:Experimental source: skin  
 C:Genetics:  
 A:Gene: Ch2

A:Introns: 21/1; 81/3  
 C:Keywords: glycoprotein  
 F:1-17/Domain: signal sequence #status predicted <SIG>  
 F:18-107/Product: major allergen chain 2, short form #status predicted <MAT>  
 F:50/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 0.6%; Score 7; DB 2; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1157 ALLVLAL 1163

Db 4 ALLVLAL 10  
 |||||

## RESULT 67

CS6413

major allergen Fel d1 chain 2 precursor - cat

C:Species: Felis silvestris catus (domestic cat)  
 C:Date: 11-Aug-1995 #sequence\_revision 11-Aug-1995 #text\_change 01-Dec-2000  
 C:Accession: CS6413; JC1145  
 R:Morgenstern, J.P.; Griffith, I.J.; Brauer, A.W.; Rogers, B.L.; Bond, J.F.; Chapman,  
 Proc. Natl. Acad. Sci. U.S.A. 88, 9690-9694, 1991  
 A:Title: Amino acid sequence of Fel d1, the major allergen of the domestic cat: prote  
 A:Reference number: A56413; MUID:92052157; PMID:1946388  
 A:Accession: CS6413

A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-109 <MOR>  
 A:Cross-references: GB:M77341; NID:g163822; PIDN:AA041616.1; PID:g163823  
 R:Griffith, I.J.; Craig, S.; Pollock, J.; Yu, X.B.; Morgenstern, J.P.; Rogers, B.L.  
 Gene 113, 263-268, 1992  
 A:Title: Expression and genomic structure of the genes encoding Fd1, the major allerg  
 A:Reference number: JC1126; MUID:92241678; PMID:1572548  
 A:Accession: JC1145

A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 18-109 <GRI>  
 A:Experimental source: salivary gland  
 C:Keywords: glycoprotein  
 F:1-17/Domain: signal sequence #status predicted <SIG>  
 F:50/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 0.6%; Score 7; DB 2; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 97;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1157 ALLVLAL 1163  
 |||||

Db 4 ALLVLAL 10

## RESULT 68

INRB

insulin precursor - rabbit

N:Alternate names: preproinsulin  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 24-Apr-1984 #sequence\_revision 23-Aug-1997 #text\_change 18-Jun-1999  
 C:Accession: A53438; A01581  
 R:Devaskar, S.U.; Giddings, S.J.; Rajakumar, P.A.; Carnaghi, L.R.; Menon, R.K.; Zahm,  
 J. Biol. Chem. 269, 8445-8454, 1994  
 A:Title: Insulin gene expression and insulin synthesis in mammalian neuronal cells.  
 A:Reference number: A53438; MUID:94179230; PMID:8132571

A:Accession: A53438  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-110 <DEV>  
 A:Cross-references: GB:U03610; NID:g467970; PIDN:AAA19033.1; PID:g467971  
 R:Smith, L.F.  
 Am. J. Med. 40, 662-666, 1966  
 A:Title: Species variation in the amino acid sequence of insulin.  
 A:Reference number: A90029; MUID:66160119; PMID:5949593

A:Accession: A01581  
 A:Molecule type: protein  
 A:Residues: 25-54; 90-110 <SMI>  
 C:Superfamily: insulin  
 C:Keywords: hormone; pancreas  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-54/Domain: insulin chain B #status experimental <BCH>  
 F:57-87/Domain: connecting C peptide #status predicted <CPBP>  
 F:90-110/Domain: insulin chain A #status experimental <ACH>  
 F:31-96; 43-109, 95-100/Disulfide bonds: #status predicted

```
Query Match      0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1155 LLALLVL 1161
Db 10 LLALLVL 16

RESULT 69
IPRT1
insulin 1 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Oct-1981 #sequence_revision 23-Oct-1991 #text_change 24-Sep-1999
A:Accession: A90788; A90789; A94231; B92120; I51945; A01589
R:Colodner, R.; Bell, G.; Tischer, E.; DeNoto, F.M.; Ullrich, A.; Pictet, R.; Rutter, W.J.
Cell 18, 533-543, 1979
A:Title: Isolation and characterization of a cloned rat insulin gene.
A:Reference number: A90788; XUID:80045034; PMID:498283
A:Accession: A90788
A:Molecule type: DNA
A:Residues: 1-110 <COR>
A:Cross-references: GB:J00747; NID:q204956; PIDN:AAA41442.1; PID:q204957
R:Lomedico, P.; Rosenthal, N.; Efstratiadis, A.; Gilbert, W.; Kolodner, R.; Tizard, R.
Cell 18, 545-558, 1979
A:Title: The structure and evolution of the two nonallelic rat preproinsulin genes.
A:Reference number: A90789; XUID:80045035; PMID:498284
A:Accession: A90789
A:Molecule type: DNA
A:Residues: 1-110 <IOM>
A:Cross-references: GB:J00747; NID:q204956; PIDN:AAA41442.1; PID:q204957
R:Steiner, D.F.; Clark, J.L.; Nollan, C.; Rubenstein, A.H.; Margolis, E.; Aten, B.; Oye
Recent Prog. Horm. Res. 25, 207-282, 1969
A:Title: Proinsulin and the biosynthesis of insulin.
A:Reference number: A94231; XUID:70067613; PMID:4311938
A:Accession: A94231
A:Molecule type: protein
A:Residues: 25-54; 90-110 <STB>
R:Tager, H.S.; Steiner, D.F.
J. Biol. Chem. 247, 7936-7940, 1972
A:Title: Primary structures of the proinsulin connecting peptides of the rat and horse.
A:Reference number: A92120; XUID:73061498; PMID:4640931
A:Accession: B92120
A:Molecule type: protein
A:Residues: 57-87 <TAG>
R:Lomedico, P.T.; Rosenthal, N.; Kolodner, R.; Efstratiadis, A.; Gilbert, W.
Ann. N.Y. Acad. Sci. 343, 425-432, 1980
A:Title: The structure of rat preproinsulin genes.
A:Reference number: I51945; XUID:80240379; PMID:6249167
A:Accession: I51945
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-110 <RES>
A:Cross-references: GB:M25584; NID:q204947; PIDN:AAA41439.1; PID:g204948
C:Genetics:
A:Gene: INS1
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <BCH>
F:55-54, 90-110/Product: insulin #status experimental <MAT>
F:57-87/Domain: connecting peptide #status experimental <CPBP>
F:90-110/Domain: insulin chain A #status experimental <ACH>
F:31-96, 43-109, 95-100/Disulfide bonds: #status experimental

Query Match      0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1155 LLALLVL 1161
Db 10 LLALLVL 16

RESULT 70
S30268
protein hdeA precursor - Escherichia coli (strain K-12)
N:Alternate names: 10K S protein precursor; hypothetical protein A
C:Species: Escherichia coli
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Mar-2002
A:Accession: S30268; S30266; S47730; A65149
R:Yoshida, T.; Ueguchi, C.; Yamada, H.; Mizuno, T.
Mol. Gen. Genet. 237, 113-122, 1993
A:Title: Function of the Escherichia coli nucleoid protein, H-NS: molecular analysis
A:Reference number: S30261; XUID:93204884; PMID:8455549
A:Accession: S30268
A:Molecule type: DNA
A:Residues: 1-110 <YOS>
A:Cross-references: EMBL:D11109; NID:g216429; PIDN:BAA01883.1; PID:g216431
A:Accession: S30266
A:Molecule type: protein
A:Residues: 22-40 <YOW>
R:Plunkett, G.
submitted to the EMBL Data Library, March 1994
A:Reference number: S47666
A:Molecule type: DNA
A:Residues: 1-110 <PLU>
A:Cross-references: EMBL:U00039; NID:q466582; PIDN:AAB18486.1; PID:q466647
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; XUID:97426617; PMID:9278503
A:Accession: A65149
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-110 <BLAT>
A:Cross-references: GB:AE000427; GB:U00096; NID:g1789919; PIDN:AAC76535.1; PID:g17899
A:Experimental source: strain K-12, substrain MGL655
C:Genetics:
A:Gene: hdeA
C:Superfamily: Escherichia coli protein hdeA
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-110/Product: 10K-S protein #status experimental <MAT>

Query Match      0.6%; Score 7; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1150 LGGILL 1156
Db 9 LGGILL 15

RESULT 71
P91177
protein hdeA precursor - Escherichia coli (strain O157:H7, substrain RMD 0509952)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 27-Nov-2001
A:Accession: P91177
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; XUID:21156231; PMID:11258796
A:Accession: P91177
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-110 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA937813.1; PID:g13363864; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECS4390
C:Superfamily: Escherichia coli protein hdeA
```

Query Match 0.6%; Score 7; DB 2; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 97;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLL 1156  
 |||||  
 Db 9 LGGLLLL 15

## RESULT 72

G86023  
 C:Species: Escherichia coli (strain O157:H7, substrain EDL933)  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 27-Nov-2001  
 C:Accession: G86023  
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: G86023  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-110 <STO>  
 A:Cross-references: GB:AE005174; NID:gl2518219; PIDN:AA058651.1; GSPDB:GN00145; UWGP:249  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: hdeA  
 C:Superfamily: Escherichia coli protein hdeA

Query Match 0.6%; Score 7; DB 2; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 97;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLL 1156  
 |||||  
 Db 9 LGGLLLL 15

## RESULT 73

C95870  
 C:Species: Sinorhizobium meliloti  
 C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
 C:Accession: C95870  
 R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo  
 A:Reference number: A95842; MUID:21396508; PMID:11481431  
 A:Accession: C95870  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-114 <XOR>  
 A:Cross-references: GB:AL591985; PIDN:CAC48627.1; PID:gl5140099; GSPDB:GN00167  
 A:Experimental source: strain 1021, megaplasmid pSymB  
 R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
 Peta, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.P.;  
 L.; Hyman, R.W.; Jones, T.  
 Science 283, 668-672, 2001  
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Voh, K.  
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A:Reference number: A96039; MUID:21368234; PMID:11474104  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: Smb20236  
 A:Genome: plasmid

Query Match 0.6%; Score 7; DB 2; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 RLGLSLA 110

Db 88 RLGLSLA 94  
 |||||

## RESULT 74

D75434  
 C:Species: Deinococcus radiodurans (strain R1)  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
 C:Accession: D75434  
 R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: D75434  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-115 <WHI>  
 A:Cross-references: GB:AE001962; GB:AE000513; NID:96458855; PIDN:AAF10702.1; PID:9645  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR1125  
 A:Map position: 1

Query Match 0.6%; Score 7; DB 2; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALL 1159  
 |||||  
 Db 10 LLLALL 16

## RESULT 75

B81215  
 C:Species: Neisseria meningitidis  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Feb-2001  
 C:Accession: B81215; S81792  
 R;Rettlein, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,  
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.  
 et al.; Qin, H.; Venter, J.C.; Gill, J.; Scarlato, V.; Masiagnani, V.; Pizza, M.  
 Science 287, 1809-1815, 2000  
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;  
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A:Reference number: A81000; MUID:20175755; PMID:10710307  
 A:Accession: B81215  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-132 <TET>  
 A:Cross-references: GB:AE002386; GB:AE002098; NID:G7225512; PIDN:AAF40742.1; PID:9722  
 A:Experimental source: serogroup B, strain MC58  
 R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mc  
 ; Holroyd, S.; Jagels, K.; Leather, S.; Moulle, S.; Mungall, K.; Quail, M.A.; Rajandre  
 Nature 404, 502-506, 2000  
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491  
 A:Reference number: A81775; MUID:20225556; PMID:10761919  
 A:Accession: G81792  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-132 <PAR>  
 A:Cross-references: GB:AL162758; GB:AL157959; NID:G7380672; PIDN:CAB85407.1; PID:9738  
 A:Experimental source: serogroup A, strain Z2491  
 C:Genetics:  
 A:Gene: NMB0291; NMA2196

QY 1151 GGLLILA 1157

Query Match 0.6%; Score 7; DB 2; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db           |||||||  
          45 GGLLIIA 51

Search completed: June 24, 2004, 18:05:04  
Job time : 32 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 17:57:31 ; Search time 18 Seconds  
(without alignments)  
3436.630 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 1188

Sequence: 1 MDLPRGLVWALSMPGFT.....FRSARRRREPLDTPPKVLE 1188

Scoring table:

OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 911   | 76.7        | 1189   | 1     | ITAH_HUMAN  |
| 2          | 15    | 1.3         | 285    | 1     | ITAI_CHICK  |
| 3          | 15    | 1.3         | 1151   | 1     | ITAI_HUMAN  |
| 4          | 15    | 1.3         | 1180   | 1     | ITAI_HUMAN  |
| 5          | 13    | 1.1         | 1167   | 1     | ITAG_HUMAN  |
| 6          | 11    | 0.9         | 1170   | 1     | ITB2_BOVIN  |
| 7          | 11    | 0.9         | 1181   | 1     | ITR2_HUMAN  |
| 8          | 9     | 0.8         | 312    | 1     | CAH4_BOVIN  |
| 9          | 9     | 0.8         | 371    | 1     | LEU3_SCHPO  |
| 10         | 9     | 0.8         | 1178   | 1     | ITR2_MOUSE  |
| 11         | 8     | 0.7         | 76     | 1     | CD24_MOUSE  |
| 12         | 8     | 0.7         | 76     | 1     | CD24_RAT    |
| 13         | 8     | 0.7         | 80     | 1     | CD24_HUMAN  |
| 14         | 8     | 0.7         | 97     | 1     | LST1_HUMAN  |
| 15         | 8     | 0.7         | 103    | 1     | FTSB_SALTY  |
| 16         | 8     | 0.7         | 105    | 1     | FTSB_KLEAE  |
| 17         | 8     | 0.7         | 179    | 1     | IL22_HUMAN  |
| 18         | 8     | 0.7         | 234    | 1     | EMB2_CAVPO  |
| 19         | 8     | 0.7         | 264    | 1     | COMT_RAT    |
| 20         | 8     | 0.7         | 402    | 1     | ORDE_PSEAE  |
| 21         | 8     | 0.7         | 423    | 1     | NSMA_HUMAN  |
| 22         | 8     | 0.7         | 444    | 1     | ENGA_MYCPE  |
| 23         | 8     | 0.7         | 451    | 1     | HISX_COREF  |
| 24         | 8     | 0.7         | 467    | 1     | D4DR_HUMAN  |
| 25         | 8     | 0.7         | 467    | 1     | SIL7_HUMAN  |
| 26         | 8     | 0.7         | 470    | 1     | CPBK_MOUSE  |
| 27         | 8     | 0.7         | 487    | 1     | XILC_PSEFU  |
| 28         | 8     | 0.7         | 491    | 1     | CPBI_RAT    |
| 29         | 8     | 0.7         | 491    | 1     | CPB2_RAT    |
| 30         | 8     | 0.7         | 500    | 1     | CPBA_MOUSE  |
| 31         | 8     | 0.7         | 604    | 1     | VE1_BPV2    |
| 32         | 8     | 0.7         | 605    | 1     | VE1_BPV1    |
| 33         | 8     | 0.7         | 811    | 1     | TMS6_HUMAN  |

|     |   |     |      |   |             |
|-----|---|-----|------|---|-------------|
| 34  | 8 | 0.7 | 1050 | 1 | ITAS_XENLA  |
| 35  | 8 | 0.7 | 1146 | 1 | ITAM_DROME  |
| 36  | 8 | 0.7 | 1152 | 1 | ITAM_HUMAN  |
| 37  | 8 | 0.7 | 1157 | 1 | DD37_HUMAN  |
| 38  | 8 | 0.7 | 1162 | 1 | ITAD_HUMAN  |
| 39  | 8 | 0.7 | 1163 | 1 | ITAD_HUMAN  |
| 40  | 8 | 0.7 | 1173 | 1 | ITAE_HUMAN  |
| 41  | 8 | 0.7 | 1179 | 1 | ITAE_HUMAN  |
| 42  | 8 | 0.7 | 1460 | 1 | PMPC_CHLMU  |
| 43  | 8 | 0.7 | 1507 | 1 | SET2_CABEL  |
| 44  | 8 | 0.7 | 1664 | 1 | RPAL_YEAST  |
| 45  | 8 | 0.7 | 2238 | 1 | RRPL_BUNYW  |
| 46  | 7 | 0.6 | 95   | 1 | LST1_MOUSE  |
| 47  | 7 | 0.6 | 109  | 1 | FEL2_FELCA  |
| 48  | 7 | 0.6 | 110  | 1 | HDEA_ECOLI  |
| 49  | 7 | 0.6 | 110  | 1 | INS1_RAT    |
| 50  | 7 | 0.6 | 110  | 1 | INS_RABIT   |
| 51  | 7 | 0.6 | 139  | 1 | IGF_MXGL    |
| 52  | 7 | 0.6 | 139  | 1 | LAMP_PETMA  |
| 53  | 7 | 0.6 | 141  | 1 | HBAO_CHICK  |
| 54  | 7 | 0.6 | 141  | 1 | HBAO_MELGA  |
| 55  | 7 | 0.6 | 142  | 1 | MK_CHICK    |
| 56  | 7 | 0.6 | 160  | 1 | BIK_HUMAN   |
| 57  | 7 | 0.6 | 160  | 1 | NUOE_AQUAE  |
| 58  | 7 | 0.6 | 161  | 1 | TCH2_ARATH  |
| 59  | 7 | 0.6 | 171  | 1 | YF87_METJA  |
| 60  | 7 | 0.6 | 176  | 1 | YD2F_SCHPO  |
| 61  | 7 | 0.6 | 194  | 1 | HIS7_CHLTE  |
| 62  | 7 | 0.6 | 201  | 1 | Z239_MOUSE  |
| 63  | 7 | 0.6 | 206  | 1 | GPBB_MOUSE  |
| 64  | 7 | 0.6 | 208  | 1 | PSMB_SULSO  |
| 65  | 7 | 0.6 | 213  | 1 | KPTA_PYPAE  |
| 66  | 7 | 0.6 | 214  | 1 | CH13_HUMAN  |
| 67  | 7 | 0.6 | 214  | 1 | CYB_CERCE   |
| 68  | 7 | 0.6 | 220  | 1 | FGF3_CHICK  |
| 69  | 7 | 0.6 | 221  | 1 | SDFL_HUMAN  |
| 70  | 7 | 0.6 | 223  | 1 | EMBP_MOUSE  |
| 71  | 7 | 0.6 | 227  | 1 | EMBP_RAT    |
| 72  | 7 | 0.6 | 233  | 1 | EMBN_CAVPO  |
| 73  | 7 | 0.6 | 235  | 1 | YBEU_ECOLI  |
| 74  | 7 | 0.6 | 235  | 1 | YOGA_ECOLI  |
| 75  | 7 | 0.6 | 236  | 1 | RR2_MAIZE   |
| 76  | 7 | 0.6 | 236  | 1 | RR2_ORYSA   |
| 77  | 7 | 0.6 | 236  | 1 | RR2_WHEAT   |
| 78  | 7 | 0.6 | 240  | 1 | IBP5_HUMAN  |
| 79  | 7 | 0.6 | 253  | 1 | GX42_RAT    |
| 80  | 7 | 0.6 | 253  | 1 | Y685_METJA  |
| 81  | 7 | 0.6 | 260  | 1 | COX3_ASTPE  |
| 82  | 7 | 0.6 | 260  | 1 | COX3_PISOC  |
| 83  | 7 | 0.6 | 261  | 1 | COX3_DINSE  |
| 84  | 7 | 0.6 | 261  | 1 | COX3_PARLI  |
| 85  | 7 | 0.6 | 262  | 1 | COX3_BRALA  |
| 86  | 7 | 0.6 | 262  | 1 | COX3_METSE  |
| 87  | 7 | 0.6 | 263  | 1 | LPSA_PASHA  |
| 88  | 7 | 0.6 | 270  | 1 | REP8_HUMAN  |
| 89  | 7 | 0.6 | 276  | 1 | NO5Y_PSEST  |
| 90  | 7 | 0.6 | 276  | 1 | UPK_EACUS   |
| 91  | 7 | 0.6 | 295  | 1 | YZ34_METJA  |
| 92  | 7 | 0.6 | 301  | 1 | PANE_PYRHO  |
| 93  | 7 | 0.6 | 306  | 1 | COAA_STRMU  |
| 94  | 7 | 0.6 | 309  | 1 | NATI_HUMAN  |
| 95  | 7 | 0.6 | 309  | 1 | NTCB_SYNP7  |
| 96  | 7 | 0.6 | 311  | 1 | MRAW_CAUCR  |
| 97  | 7 | 0.6 | 311  | 1 | PYDA_ENTFA  |
| 98  | 7 | 0.6 | 312  | 1 | O8H2_HUMAN  |
| 99  | 7 | 0.6 | 313  | 1 | LEU3_NELIA  |
| 100 | 7 | 0.6 | 315  | 1 | CALU1_HUMAN |

#### ALIGNMENTS

RESULT 1

ITAH\_HUMAN STANDARD; PRT; 1189 AA.  
AC Q9UKX5; Q9URQ1;  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-11 precursor.  
GN ITGAL1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal heart, and Osteoblast;  
RX MEDLINE=99417679; PubMed=10486209;  
RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,  
RA Wang S.-X., Morris C.M., Krissansen G.W.;  
RT "Cloning, sequence analysis, and chromosomal localization of the novel  
RL human integrin alpha11 subunit (ITGAL1).";  
RL Genomics 60:179-187(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal muscle, and Uterus;  
RX MEDLINE=99395147; PubMed=10464311;  
RA Velling T., Kusche-Gullberg M., Seijersen T., Gullberg D.;  
RT "cDNA Cloning and Chromosomal Localization of Human alpha(11)  
RT Integrin. A collagen-binding, i domain-containing, beta(1)-associated  
RT integrin alpha-chain present in muscle tissues.";  
RL J. Biol. Chem. 274:25735-25742(1999).  
RN [3]  
RP SEQUENCE OF 954-1188 FROM N.A.  
RC TISSUE=Fibroblast;  
RA Andreu N., Estvill X., Escarceller M., Sumoy L.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11  
CC ASSOCIATES WITH BETA-1.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND  
CC HEART, INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO  
CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO  
CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS, IN  
CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN  
CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.  
CC -!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING  
CC FETAL MUSCLE CELLS (IN VITRO).  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; AF109681; AAF01258.1; --  
DR GO; GO:0008305; C:integrin complex; TAS.  
DR EMBL; AF137378; AAD51919.2; --  
DR EMBL; AL358064; CAB94392.1; --  
DR HSSP; P17301; IAOX.  
DR Genew; HGNC:6136; ITGAL1.  
DR MIM; 604789; --  
DR GO; GO:0008305; C:integrin complex; TAS.  
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.  
DR GO; GO:0005518; F:collagen binding; TAS.  
DR GO; GO:0007160; P:cell-matrix adhesion; TAS.  
DR GO; GO:0007517; P:muscle development; TAS.  
DR InterPro; IPR000413; Integrin\_alpha.

DR InterPro; IPR002035; VWFA\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR00453; VWFADOMAIN.  
DR SMART; SM00131; Int\_alpha; 5.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; FALSE\_NEG.  
DR PROSITE; PS50234; VWFA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Signal; Repeat; Calcium; Magnesium; Polymorphism.  
FT SIGNAL 1 22  
FT CHAIN 23 1189  
FT DOMAIN 23 1142  
FT TRANSMEM 1143 1165  
FT DOMAIN 1166 1189  
FT REPEAT 38 94  
FT REPEAT 102 163  
FT DOMAIN 167 345  
FT REPEAT 359 420  
FT REPEAT 422 475  
FT REPEAT 477 537  
FT REPEAT 539 598  
FT REPEAT 601 653  
FT DOMAIN 1154 1162  
FT DOMAIN 1174 1177  
FT CA\_BIND 488 496  
FT CA\_BIND 551 559  
FT CA\_BIND 613 621  
FT DISULFID 76 83  
FT DISULFID 121 139  
FT DISULFID 129 159  
FT DISULFID 659 668  
FT DISULFID 674 729  
FT DISULFID 781 787  
FT DISULFID 881 893  
FT CARBOHYD 82 82  
FT CARBOHYD 95 95  
FT CARBOHYD 291 291  
FT CARBOHYD 331 331  
FT CARBOHYD 358 358  
FT CARBOHYD 449 449  
FT CARBOHYD 462 462  
FT CARBOHYD 528 528  
FT CARBOHYD 642 642  
FT CARBOHYD 694 694  
FT CARBOHYD 857 857  
FT CARBOHYD 894 894  
FT CARBOHYD 973 973  
FT CARBOHYD 1032 1032  
FT CARBOHYD 1040 1040  
FT VARIANT 433 433  
FT VARIANT 524 524  
FT VARIANT 972 972  
FT VARIANT 1003 1003  
FT VARIANT 1030 1030  
FT VARIANT 1094 1094  
FT SEQUENCE 1189 AA; 133609 MW; 60303C08A4A4CD52 CRC64;  
Query Match 76.7%; Score 911; DB 1; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDLPRGLVWAWALSLWPCFTDTFMNDRKPRVTPGSRTPFGYTVQQHDSGNKWLVVGA 60  
Db 1 MDLPRGLVWAWALSLWPCFTDTFMNDRKPRVTPGSRTPFGYTVQQHDSGNKWLVVGA 60

QY 61 PLENGYQKTDVYKCPVHGNCTKLNIGRVTLNSVSRKDNVRLGLSLATNPKNISFLA 120  
 DB 61 PLENGYQKTDVYKCPVHGNCTKLNIGRVTLNSVSRKDNVRLGLSLATNPKNISFLA 120  
 QY 121 CSPLMSHSCGSSYYTTCGSRVNSFRSKTVAPALQCCOTYMDIVIVLDGNSIYPAVE 180  
 DB 121 CSPLMSHSCGSSYYTTCGSRVNSFRSKTVAPALQCCOTYMDIVIVLDGNSIYPAVE 180  
 QY 181 VQFLINILKFFYIGFGQIQGVVQYGEDVVEHFLNDRYSVKDVVEAASHIEQGGTET 240  
 DB 181 VQFLINILKFFYIGFGQIQGVVQYGEDVVEHFLNDRYSVKDVVEAASHIEQGGTET 240  
 QY 241 RTAGIEFARAEAFQKGRKAKKVMIVITGESHDSPLDKVYQSSRDVNTYAVAVL 300  
 DB 241 RTAGIEFARAEAFQKGRKAKKVMIVITGESHDSPLDKVYQSSRDVNTYAVAVL 300  
 QY 301 GYNNRGINPFTFNEIKYIASDDPKHFFNVDEAALKDVIDALGRIFSLGNTKNET 360  
 DB 301 GYNNRGINPFTFNEIKYIASDDPKHFFNVDEAALKDVIDALGRIFSLGNTKNET 360  
 QY 361 SFGLMSOTGFSHVVDGVLGAVGAYDMGAVLKTSAGKVIPLRESYKKEPBEILKN 420  
 DB 361 SFGLMSOTGFSHVVDGVLGAVGAYDMGAVLKTSAGKVIPLRESYKKEPBEILKN 420  
 QY 421 HGAVLGYTVTSVSSRQGRVYVAGAPRNHTGKVLFTMNNRSLTHQAMRGQIQISYP 480  
 DB 421 HGAVLGYTVTSVSSRQGRVYVAGAPRNHTGKVLFTMNNRSLTHQAMRGQIQISYP 480  
 QY 481 GSEITSVDIDGGVTDVLLGAPMYFNEGRGKVVYVELQNRFPVYNGTLKDSHSYQNA 540  
 DB 481 GSEITSVDIDGGVTDVLLGAPMYFNEGRGKVVYVELQNRFPVYNGTLKDSHSYQNA 540  
 QY 541 RFGSSIASVRDLNDSYNDVVGAPLEDNHAGAIYFHGPGSLKTPKQBITASELATG 600  
 DB 541 RFGSSIASVRDLNDSYNDVVGAPLEDNHAGAIYFHGPGSLKTPKQBITASELATG 600  
 QY 601 LQYFGCSIHGQDLNEDGLIDLAVGALNAVILSRPVVQINASLHFEPSKINIFHRDCK 660  
 DB 601 LQYFGCSIHGQDLNEDGLIDLAVGALNAVILSRPVVQINASLHFEPSKINIFHRDCK 660  
 QY 661 RSGRDATCLAFCTPIPLAPHQTTTGVRYNATMDERYTTPRAHLDEGGDFTNRVAV 720  
 DB 661 RSGRDATCLAFCTPIPLAPHQTTTGVRYNATMDERYTTPRAHLDEGGDFTNRVAV 720  
 QY 721 LLSGQELCERINFHVLDTADYVRPVTSEYSLDPDGHGMLDGGWPTTLRVSVFPWNG 780  
 DB 721 LLSGQELCERINFHVLDTADYVRPVTSEYSLDPDGHGMLDGGWPTTLRVSVFPWNG 780  
 QY 781 CNEDEHCVDPDLVDARSOLPTAMEYCORVLRKPQDCSAYTSLSDTTVFIIESTQRVAV 840  
 DB 781 CNEDEHCVDPDLVDARSOLPTAMEYCORVLRKPQDCSAYTSLSDTTVFIIESTQRVAV 840  
 QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKESDGSIECVNEERLQKQVCNVSYPFF 900  
 DB 841 EATLENRGENAYSTVLNISQSANLQFASLIQKESDGSIECVNEERLQKQVCNVSYPFF 900  
 QY 901 RAKAKVAFRLD 911  
 DB 901 RAKAKVAFRLD 911

RESULT 2

ITAL CHICK STANDARD; PRT; 285 AA.  
 AC Q90615;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Integrin alpha-1 (laminin and collagen receptor) (VLA-1) (CD49a)  
 DE (fragment).  
 GN ITGAL.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 ON NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Gizzard;  
 RX MEDLINE=94357930; PubMed=7521332;  
 RA Kern A., Brisevitz R., Bank I., Marcantonio E.E.;  
 RT "The role of the I domain in ligand binding of the human integrin  
 alpha 1 beta 1.";  
 RL J Biol. Chem. 269:22811-22816 (1994).  
 CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND  
 COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-  
 Z-R IN COLLAGEN.  
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1  
 ASSOCIATES WITH BETA-1.  
 CC -!- SUBCELLULAR LOCATION: TYPE I membrane protein.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; U10114; AAA59067.1; -.  
 DR HSP; P17301; IAOX.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWFA.  
 DR Pfam; PF00092; vwa; 1.  
 DR SMART; SM00327; VWFA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; PARTIAL.  
 DR PROSITE; PS00234; VWFA; 1.  
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane.  
 PT NON TER 1 1  
 FT DOMAIN <1 285 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 66 279 VWFA.  
 FT CARBOHYD 2 2 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT NON TER 285 285  
 SQ SEQUENCE 285 AA; 31503 MW; 1B05D3246CD5C47E CRC64;  
 Query Match 1.3%; Score 15; DB 1; Length 285;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 164 DIVIVLDGNSIYPM 178  
 DB 63 DIVIVLDGNSIYPM 77  
 RESULT 3  
 ITAL HUMAN STANDARD; PRT; 1151 AA.  
 ID -ITAL HUMAN  
 AC P56139;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DB Integrin alpha-1 (laminin and collagen receptor) (VLA-1) (CD49a).  
 GN ITGAL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.



```

DR EMBL; X52140; CAA36384.1; -.
DR PIR; A35854; A35854.
DR PDB; 1CR4; 03-MAY-00.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Calcium; Magnesium; 3D-structure.
FT SIGNAL 1 28
FT CHAIN 29 1180 INTEGRIN_ALPHA-1.
FT DOMAIN 29 1142 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1143 1165 POTENTIAL.
FT DOMAIN 1166 1180 CYTOPLASMIC (POTENTIAL).
FT REPEAT 44 103 FG-GAP 1.
FT REPEAT 2 2 FG-GAP 2.
FT DOMAIN 175 388 VWFA.
FT REPEAT 377 432 FG-GAP 3.
FT REPEAT 433 484 FG-GAP 4.
FT REPEAT 485 565 FG-GAP 5.
FT REPEAT 567 626 FG-GAP 6.
FT REPEAT 629 681 FG-GAP 7.
FT CA_BIND 497 505 POTENTIAL.
FT CA_BIND 579 587 POTENTIAL.
FT CA_BIND 641 649 POTENTIAL.
FT SITE 1168 1172 GFPR MOTIF.
FT DISULFID 82 92 BY SIMILARITY.
FT DISULFID 687 696 BY SIMILARITY.
FT DISULFID 702 755 BY SIMILARITY.
FT DISULFID 807 813 BY SIMILARITY.
FT DISULFID 877 885 BY SIMILARITY.
FT DISULFID 1029 1082 BY SIMILARITY.
FT DISULFID 1066 1073 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 317 317 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 341 341 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 698 698 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 747 747 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 779 779 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 820 820 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 839 839 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 882 882 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 907 907 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 965 965 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 973 973 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 1007 1007 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 1084 1084 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 1103 1103 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 1114 1114 N-LINKED (GLCNAC. ) (POTENTIAL).
FT SEQUENCE 1180 AA; 130808 MW; 8B5DA2B802362EB4 CRC64;

```

Query Match 1.3%; Score 15; DB 1; Length 1180;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 164 DIVIVLDGNSIYPW 178

DB 172 DIVIVLDGNSIYPW 186

```

RESULT 5
ITAG HUMAN
ID ITAG_HUMAN STANDARD; PRT; 1167 AA.
AC O75578; Q9URZ8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-10 precursor.
GN ITGA10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_taxid=9606;
EN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Articular chondrocytes;
RX MEDLINE=98352078; PubMed=9685391;
RA Camper L., Hellman U., Lundgren-Akerlund E.;
RT "Isolation, cloning, and sequence analysis of the integrin subunit
alpha10, a beta1-associated collagen binding integrin expressed on
chondrocytes.";
RT J. Biol. Chem. 273:20383-20389(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Endothelial cells, and Heart;
RX MEDLINE=20169197; PubMed=10702680;
RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,
Wang S.-X., Langley R., Krissansen G.W.;
RT "The integrin alpha10 subunit: expression pattern, partial gene
structure, and chromosomal localization.";
RT Cytogenet. Cell Genet. 87:238-244(1999).
CC -!- FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Widely expressed with highest expression in
muscle and heart. Found in articular cartilage.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF074015; AAC31952.1; -.
EMBL; AF112345; AAF21944.1; -.
EMBL; AF172723; AAF61638.1; -.
HSSP; P17301; LAOX.
Genew; HGNC:6135; ITGA10.
MIM; 604042; -.
GO; GO:0008305; C:integrin complex; TAS.
GO; GO:0004895; F:cell adhesion receptor activity; TAS.
GO; GO:0005518; F:collagen binding; TAS.
GO; GO:0007160; P:cell-matrix adhesion; TAS.
InterPro; IPR000413; Integrin_alpha.
InterPro; IPR02035; VWFA.
Pfam; PF01839; FG-GAP; 3.
Pfam; PF00092; vwa; 1.
PRINTS; PR01185; INTEGRINA.
PRINTS; PR00453; VWFADOMAIN.
SMART; SM00191; Int_alpha; 4.
SMART; SM00327; VWA; 1.
PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.
PROSITE; PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Signal; Repeat; Calcium; Magnesium.

```

FT SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 1167 INTEGRIN ALPHA-10.  
 FT DOMAIN 23 1122 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1123 1145 POTENTIAL.  
 FT DOMAIN 1146 1167 CYTOPLASMIC (POTENTIAL).  
 FT REPEAT 38 97 FG-GAP 1.  
 FT REPEAT 7 7 FG-GAP 2.  
 FT DOMAIN 167 350 VWFA.  
 FT REPEAT 365 427 FG-GAP 3.  
 FT REPEAT 428 482 FG-GAP 4.  
 FT REPEAT 546 605 FG-GAP 5.  
 FT REPEAT 608 660 FG-GAP 6.  
 FT DOMAIN 1134 1140 POLY-LEU.  
 FT CA\_BIND 494 502 POTENTIAL.  
 FT CA\_BIND 558 566 POTENTIAL.  
 FT CA\_BIND 620 628 POTENTIAL.  
 FT DISULFID 76 86 BY SIMILARITY.  
 FT DISULFID 666 675 BY SIMILARITY.  
 FT DISULFID 681 736 BY SIMILARITY.  
 FT DISULFID 789 795 BY SIMILARITY.  
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 921 921 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 844 844 I -> L (IN REF. 2).  
 FT CONFLICT 909 909 G -> V (IN REF. 2).  
 FT CONFLICT 926 926 E -> D (IN REF. 2).  
 SQ SEQUENCE 1167 AA; 127573 MW; A57D3A1C25C1AAE0 CRC64;

Query Match 1.1%; Score 13; DB 1; Length 1167;  
 Best Local Similarity 100.0%; Pred. No. 0.00054;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 VIVLDGNSIYPW 178  
 |||||  
 DB 169 VIVLDGNSIYPW 181

RESULT 6  
 ID\_ITA2\_BOVIN STANDARD; PRT; 1170 AA.  
 AC P53710;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DE 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GPIa)  
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).  
 GN ITGA2.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94193647; PubMed=75111592;  
 RA Kamata T., Puzon W., Takada Y.;  
 RT "Identification of putative ligand binding sites within I domain of  
 NT integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";  
 RL J. Biol. Chem. 269:19659-19663(1994).  
 CC -1- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,  
 CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND B-CADHERIN. IT  
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-E-R IN  
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER  
 CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE

CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
 CC EXTRACELLULAR MATRIX.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2  
 CC ASSOCIATES WITH BETA-1.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -1- SIMILARITY: Contains 1 VWFA domain.  
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or  
 CC send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; L25886; AAB59255.1; -;  
 DR PIR; I45914; I45914.  
 DR HSSP; P17301; LAOX.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWF\_A.  
 DR Pfam; PF01839; FG-GAP; 3\_A.  
 DR Pfam; PF00357; Integrin\_A; 1.  
 DR Pfam; PF00092; vwa; 1.  
 DR SMART; SM00191; Int\_alpha; 5.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS50234; VWFA; 1.  
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 18 POTENTIAL.  
 FT CHAIN 19 1170 INTEGRIN ALPHA-2.  
 FT DOMAIN 19 1121 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1122 1143 POTENTIAL.  
 FT DOMAIN 1144 1170 CYTOPLASMIC (POTENTIAL).  
 FT REPEAT 34 92 FG-GAP 1.  
 FT REPEAT 7 7 FG-GAP 2.  
 FT DOMAIN 177 367 VWFA.  
 FT REPEAT 7 7 FG-GAP 3.  
 FT REPEAT 423 475 FG-GAP 4.  
 FT REPEAT 477 538 FG-GAP 5.  
 FT REPEAT 540 599 FG-GAP 6.  
 FT REPEAT 604 656 FG-GAP 7.  
 FT CA\_BIND 488 496 POTENTIAL.  
 FT CA\_BIND 552 560 POTENTIAL.  
 FT CA\_BIND 616 624 POTENTIAL.  
 FT SITE 472 474 CELL ATTACHMENT SITE (POTENTIAL).  
 FT SITE 1146 1150 GPFXR MOTIF.  
 FT DISULFID 72 81 BY SIMILARITY.  
 FT DISULFID 669 726 BY SIMILARITY.  
 FT DISULFID 778 784 BY SIMILARITY.  
 FT DISULFID 854 865 BY SIMILARITY.  
 FT DISULFID 1008 1039 BY SIMILARITY.  
 FT DISULFID 1044 1049 BY SIMILARITY.  
 FT CARBOHYD 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 449 449 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 688 688 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 748 748 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 945 945 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1063 1063 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1070 1070 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARIANT 580 580 G -> V.  
 FT VARIANT 588 588 R -> K.  
 FT VARIANT 725 725 R -> S.  
 SQ SEQUENCE 1170 AA; 128929 MW; BECEFF1C5F2448FB1 CRC64;

Query Match 0.9%; Score 11; DB 1; Length 1170;  
 Best Local Similarity 100.0%; Pred. No. 0.05;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVNGAPMY 505  
 |||||  
 DB 495 TDVLLVNGAPMY 505

RESULT 7

IT22 HUMAN STANDARD; PRT; 1181 AA.

AC P17301;  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GPIa)  
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).  
 GN ITGA2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.  
 RC TISSUE-Endothelial cells;  
 RX MEDLINE=89308879; PubMed=2545729;  
 RA Takada Y., Hemler M.E.;  
 RT "The primary structure of the VLA-2/collagen receptor alpha 2 subunit  
 (platelet GPIa): homology to other integrins and the presence of a  
 RT possible collagen-binding domain.";  
 RL J. Cell Biol. 109:397-407(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,  
 RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBAJ databases.  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.  
 RX MEDLINE=98019223; PubMed=9353312;  
 RA Emley J., King S.L., Bergelson J.M., Liddington R.C.;  
 RT "Crystal structure of the I domain from integrin alpha2beta1.";  
 RL J. Biol. Chem. 272:28512-28517(1997).  
 RN [4]  
 RP VARIANT HPA-5 (BR).  
 RX MEDLINE=94043762; PubMed=7901236;  
 RA Santoso S., Kalb R., Waika M., Kiefel V., Mueller-Eckhardt C.,  
 RA Newman P.J.;  
 RT "The human platelet alloantigens Br(a) and Br(b) are associated with a  
 RT single amino acid polymorphism on glycoprotein Ia (integrin subunit  
 RT alpha 2).";  
 RL J. Clin. Invest. 92:2427-2432(1993).  
 RN [5]  
 RP VARIANT GLU-534.  
 RX MEDLINE=20206009; PubMed=10741412;  
 RA Kroll H., Gardemann A., Fechter A., Haberbusch W., Santoso S.;  
 RT "The impact of the glycoprotein Ia collagen receptor subunit A1648G  
 RT gene polymorphism on coronary artery disease and acute myocardial  
 RT infarction.";  
 RL Thromb. Haemost. 83:392-395(2000).  
 CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,  
 CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT  
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN  
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER  
 CC CELLS TO COLLAGENS. MODULATION OF COLLAGEN AND COLLAGENASE GENE  
 CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
 CC EXTRACELLULAR MATRIX.  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2  
 CC associates with beta-1. Interacts with HPSS.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

-!- POLYMORPHISM: Position 534 is associated with platelet-specific  
 alloantigen HPA-5 (BR). HPA-5A/BR(A) has Lys-534 and HPA-5B/BR(B)  
 has Glu-534. HPA-5B is involved in neonatal alloimmune  
 thrombocytopenia (NAIT or NATP). The K534E polymorphism may play a  
 role in coronary artery disease (CAD).  
 -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 -!- SIMILARITY: Contains 1 VWFA domain.  
 -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 -!- DATABASE: NAME=PRO; NOTE=CD guide CD49b entry;  
 WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49b.htm".

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

EMBL; X17033; CAA34894.1; --  
 EMBL; AF512556; AAM34795.1; --  
 PIR; A33998; A33998.  
 PDB; 1A0X; 25-NOV-98.  
 PDB; 1DZ1; 02-AUG-01.  
 Genew; HGNC:6137; ITGA2.  
 MIM; 192374; --  
 DR GO; 0008305; C.integrin complex; TAS.  
 DR GO; 0005886; C.plasma membrane; TAS.  
 DR GO; 0004895; F.cell adhesion receptor activity; TAS.  
 DR GO; 0005518; F.collagen binding; TAS.  
 DR GO; 0007596; P.blood coagulation; TAS.  
 DR GO; 0007160; P.cell-matrix adhesion; TAS.  
 DR GO; 0007397; P.histogenesis and organogenesis; TAS.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWFA.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00357; integrin\_A; 1.  
 DR Pfam; PF00092; vwa; 1.  
 DR SMART; SM00191; Int\_alpha; 5.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS00234; VWFA; 1.  
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;  
 3D-structure.

|             |      |      |                                     |
|-------------|------|------|-------------------------------------|
| FT SIGNAL   | 1    | 29   | INTEGRIN ALPHA-2                    |
| FT CHAIN    | 30   | 1181 | EXTRACELLULAR (POTENTIAL).          |
| FT DOMAIN   | 1133 | 1154 | POTENTIAL.                          |
| FT TRANSMEM | 1155 | 1181 | CYTOPLASMIC (POTENTIAL).            |
| FT DOMAIN   | 1155 | 1161 | INTERACTION WITH HPSS.              |
| FT REPEAT   | 45   | 103  | FG-GAP 1.                           |
| FT REPEAT   | ?    | ?    | FG-GAP 2.                           |
| FT DOMAIN   | 188  | 378  | VWFA.                               |
| FT REPEAT   | 378  | 433  | FG-GAP 3.                           |
| FT REPEAT   | 434  | 486  | FG-GAP 4.                           |
| FT REPEAT   | 488  | 549  | FG-GAP 5.                           |
| FT REPEAT   | 551  | 610  | FG-GAP 6.                           |
| FT REPEAT   | 615  | 667  | FG-GAP 7.                           |
| FT CA_BIND  | 499  | 507  | POTENTIAL.                          |
| FT CA_BIND  | 563  | 571  | POTENTIAL.                          |
| FT CA_BIND  | 627  | 635  | POTENTIAL.                          |
| FT SITE     | 1157 | 1161 | GPFR MOTIF.                         |
| FT DISULFID | 83   | 92   | BY SIMILARITY.                      |
| FT DISULFID | 680  | 737  | BY SIMILARITY.                      |
| FT DISULFID | 789  | 795  | BY SIMILARITY.                      |
| FT DISULFID | 865  | 876  | BY SIMILARITY.                      |
| FT DISULFID | 1019 | 1050 | BY SIMILARITY.                      |
| FT DISULFID | 1055 | 1060 | BY SIMILARITY.                      |
| FT CARBOHYD | 105  | 105  | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT CARBOHYD | 112  | 112  | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT CARBOHYD | 343  | 343  | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT CARBOHYD | 432  | 432  | N-LINKED (GLCNAC. . .) (POTENTIAL). |



```

FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 693 693 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1057 1057 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1081 1081 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT VARIANT 534 534 K -> E (IN ALLOANTIGEN HPA-59;
  GSNP:1801106).
  /FTid=VAR_003977.
FT TURN 170 171
FT STRAND 173 180
FT TURN 183 184
FT TURN 188 199
FT TURN 200 201
FT STRAND 204 204
FT TURN 206 207
FT STRAND 209 216
FT STRAND 220 224
FT TURN 226 228
FT TURN 232 240
FT TURN 241 241
FT TURN 242 262
FT TURN 263 264
FT TURN 265 268
FT TURN 269 269
FT STRAND 275 282
FT TURN 289 291
FT TURN 292 301
FT TURN 302 303
FT TURN 304 311
FT TURN 313 317
FT TURN 318 319
FT TURN 323 330
FT TURN 331 332
FT TURN 337 340
FT TURN 341 344
FT TURN 347 353
FT TURN 354 362
FT TURN 363 363
SQ SEQUENCE 1181 AA, 129295 MW; 7E1B7ED968A94070 CRC64;

Query Match 0.9%; Score 11; DB 1; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAPMY 505
Db 506 TDVLLVGAPMY 516
|||||

RESULT 8
CAH4 BOVIN STANDARD; PRT; 312 AA.
AC Q95323;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Carbonic anhydrase IV precursor (EC 4.2.1.1) (Carbonate dehydratase IV) (CA-IV).
GN CA4.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Holstein; TISSUE=Kidney;
RA Tanai S.;
CC Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Reversible hydration of carbon dioxide.
CC -!- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
CC -!- COFACTOR: Zinc.

```

```

CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
  family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL; U58870; AAB09466.1; -.
CC HSSP; P22748; 1ZNC.
DR InterPro; IPR001148; Euk_Coanhd.
DR Pfam; PF00194; carb_anhydrase; 1.
DR ProDom; PD000865; Euk_Coanhd; 1.
DR PROSITE; PS00162; EUK_CO2_ANHYDRASE; FALSE NEG.
KW Glycoprotein; Lipoprotein; GPI-anchor; Membrane; Lyase; Zinc; Signal.
FT SIGNAL 1 18
FT CHAIN 19 284
FT PROPEP 285 312
FT METAL 115 115
FT METAL 117 117
FT METAL 140 140
FT DISULFID 24 36
FT DISULFID 46 229
FT LIPID 284
FT SEQUENCE 312 AA; 35151 MW; BAE320C09426351 CRC64;

Query Match 0.8%; Score 9; DB 1; Length 312;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVLA 1162
Db 3 LLLALLVLA 11
|||||

RESULT 9
LEU3_SCHPO STANDARD; PRT; 371 AA.
AC P18869;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 3-isopropylmalate dehydrogenase (EC 1.1.1.85) (Beta-IPM dehydrogenase)
DE (IMDH) (3-IPM-DH).
GN LEU1 OR SPEC1A4.02C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89106270; PubMed=3063400;
RA Kikuchi Y., Kitazawa Y., Shimatake G., Yamamoto M.;
RT "The primary structure of the leu1+ gene of Schizosaccharomyces
  pombe".
RL Curr. Genet. 14:375-379(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Guillermo R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Gobie A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

```

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volktaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Woestl D., Hilbert H.,  
 RA Borycz K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purrelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen O., Fotschkin J.,  
 RA Shpakowski G.V., Usery D., Barrell B.G., Nurse P.,  
 RA The genome sequence of *Schizosaccharomyces pombe*.  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: Catalyzes the oxidation of 3-carboxy-2-hydroxy-4-  
 CC methylpentanoate (3-isopropylmalate) to 3-carboxy-4-methyl-2-  
 CC oxopentanoate. The product decarboxylates to 4-methyl-2-  
 CC oxopentanoate.  
 CC -!- CATALYTIC ACTIVITY: 3-carboxy-2-hydroxy-4-methylpentanoate +  
 CC NAD(+) = 3-carboxy-4-methyl-2-oxopentanoate + NADH.  
 CC -!- PATHWAY: Leucine biosynthesis; third step.  
 CC -!- SUBUNIT: Homodimer (by similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to the isocitrate and isopropylmalate  
 CC dehydrogenases family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M36910; AAA35316.1; -.  
 CC EMBL; AL031174; CAA20106.1; -.  
 CC PIR; T43407; T43407.  
 CC HSSP; P12010; 2AVO.  
 CC GeneDB\_Spombe; SPEC1A4.02c; -.  
 CC InterPro; IPR001804; IsoDH.  
 CC InterPro; IPR004429; LeuB.  
 CC Pfam; PF00180; IsoDH; 1.  
 CC TIGRFAMs; TIGR00169; leuB; 1.  
 CC PROSITE; PS00470; IDH IMDH; 1.  
 CC Oxidoreductase; Leucine biosynthesis; NAD.  
 KW SEQUENCE 371 AA; 39732 MW; 65AA2B6AA94D45EE CRC64;  
 Query Match 0.8%; Score 9; DB 1; Length 371;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 378 DGVLLGAVG 386  
 DB 69 DGVLLGAVG 77  
 RESULT 10  
 ID ITA2 MOUSE  
 AC Q62469; Q62163; PRT; 1178 AA.  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GPIa)  
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).  
 GN ITGA2.  
 GN Itga2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CS7BL/6 X CBA; TISSUE=Lung;  
 RX MEDLINE=94363406; PubMed=8081889;  
 RA Edelman J.M., Chan B.M., Unival S., Onodera H., Wang D.Z.,  
 RA Damjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.,  
 RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but  
 RL not virus binding.";  
 RL Cell Adhes. Commun. 2:131-143(1994).  
 [2]  
 RP SEQUENCE OF 450-1178 FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=94355691; PubMed=7521231;  
 RA Wu J.B., Santoro S.A.,  
 RT "Complex patterns of expression suggest extensive roles for the alpha  
 RL 2 beta 1 integrin in murine development.";  
 RL Dev. Dyn. 199:292-314(1994).  
 CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING  
 CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO  
 CC COLLAGENS. MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,  
 CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
 CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN  
 CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION  
 CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.  
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2  
 CC ASSOCIATES WITH BETA-1.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Z29987; CAA82877.1; -.  
 CC EMBL; X75427; CAA53178.1; -.  
 CC PIR; S44142; S44142.  
 CC HSSP; P17301; IAOX.  
 CC MSD; MGI:96600; Itga2.  
 CC InterPro; IPR000413; Integrin\_alpha.  
 CC InterPro; IPR002035; VWF\_A.  
 CC Pfam; PF01839; FG-GAP; 3\_  
 CC Pfam; PF00357; integrin\_A; 1.  
 CC SMART; SM00191; Int\_alpha; 4.  
 CC SMART; SM00327; VWFA; 1.  
 CC PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 CC PROSITE; PS02334; VWFA; 1.  
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Platelet; Signal; Repeat; Calcium; Magnesium.  
 FT SIGNAL 1 26  
 FT CHAIN 27 1178  
 FT DOMAIN 27 1129  
 FT TRANSMEM 1130 1151  
 FT DOMAIN 1152 1178  
 FT REPEAT 42 100  
 FT REPEAT ? ?  
 FT REPEAT 185 375  
 FT DOMAIN ? ?  
 FT REPEAT 431 483  
 FT REPEAT 485 546  
 FT REPEAT 548 607  
 FT REPEAT 612 664  
 FT CA\_BIND 496 504  
 FT POTENTIAL.  
 FT CYTOPLASMIC (POTENTIAL).  
 FT EG-GAP 1.  
 FT EG-GAP 2.  
 FT VWFA.  
 FT FG-GAP 3.  
 FT FG-GAP 4.  
 FT FG-GAP 5.  
 FT FG-GAP 6.  
 FT FG-GAP 7.  
 FT POTENTIAL.

```

FT CA BIND 560 568 POTENTIAL.
FT CA_BIND 624 632 POTENTIAL.
FT SITE 480 482 CBLL ATTACHMENT SITE (POTENTIAL).
FT SITE 1154 1158 GPKR MOTIF.
FT DISULFID 80 89 BY SIMILARITY.
FT DISULFID 677 734 BY SIMILARITY.
FT DISULFID 786 792 BY SIMILARITY.
FT DISULFID 862 873 BY SIMILARITY.
FT DISULFID 1016 1047 BY SIMILARITY.
FT DISULFID 1052 1057 BY SIMILARITY.
FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 429 429 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 472 472 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 696 696 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1054 1054 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1178 AA; 128926 MW; 1P194B9C0240F465 CRC64;

Query Match 0.8%; Score 9; DB 1; Length 1178;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGP 503
Db 503 TDVLLVGP 511

RESULT 11
CD24_MOUSE
ID CD24_MOUSE STANDARD; PRT; 76 AA.
AC P24807; P26691.
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Signal transducer CD24 precursor (M1/69-J11D heat stable antigen)
DE (HSA) (Nectadrin) (LY-52) (X82 heat stable antigen) (R13-AG).
GN CD24 OR CD24A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
(1)
RN SEQUENCE FROM N.A.
RX MEDLINE=90361906; PubMed=2118158;
RA Kay R., Takei F., Humphries R.K.;
RT "Expression cloning of a cDNA encoding M1/69-J11D heat-stable
antigens."
RL J. Immunol. 145:1952-1959(1990).
(2)
RN SEQUENCE FROM N.A.
RX STRAIN=C57BL/6 X CBA, and Swiss albino X BALB/c; TISSUE=Spleen;
RA MEDLINE=91209380; PubMed=2019286;
RA Wenger R.H., Ayane M., Bose R., Koehler G., Nielsen P.J.;
RT "The genes for a mouse hematopoietic differentiation marker called
the heat-stable antigen."
RL Eur. J. Immunol. 21:1039-1046(1991).
(3)
RN REVISIONS.
RA Nielsen P.J.;
RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
(4)
RN SEQUENCE FROM N.A.
RX STRAIN=C57BL/6 X CBA; TISSUE=Spleen;
RA MEDLINE=94043127; PubMed=8226859;
RA Wenger R.H., Rochelle J.M., Seldin M.F., Koehler G., Nielsen P.J.;
RT "The heat stable antigen (mouse CD24) gene is differentially
regulated but has a housekeeping promoter."
RL J. Biol. Chem. 268:23345-23352(1993).
(5)
RN SEQUENCE OF 27-53.

```

```

RC STRAIN=C57BL/6;
RX MEDLINE=92412120; PubMed=1530634;
RA Hitsumoto Y., Nakano A., Ohnishi H., Hamada F., Sabeki S.,
RA Takeuchi N.;
RT "Purification of the murine heat-stable antigen from erythrocytes."
RL Biochem. Biophys. Res. Commun. 187:773-777(1992).
CC -!- FUNCTION: May have a specific role to play in early thymocyte
development.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- TISSUE SPECIFICITY: In lymphoid, myeloid, and erythroid cells.
CC -!- PTM: Extensively O-glycosylated (By similarity).
CC -!- SIMILARITY: TO OTHER MAMMALIAN SPECIES CD24.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; M58661; AAA39481.1; -
DR EMBL; X56469; CAA39841.1; -
DR EMBL; X72910; CAA51415.1; -
DR EMBL; X53825; CAA37822.1; -
DR PIR; A43537; A43537
DR MGD; MGI:88323; Cd24a.
KW Glycoprotein; GPI-anchor; Membrane; Signal; Antigen; Lipoprotein.
FT SIGNAL 1 26
FT CHAIN 27 53 SIGNAL TRANSDUCER CD24.
FT PROPEP 54 76 REMOVED IN MATURE FORM (POTENTIAL).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .).
FT CARBOHYD 30 30 O-LINKED (PROBABLE).
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .).
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .).
FT LIPID 53 53 GPI-anchor amidated glycine (Potential).
SQ SEQUENCE 76 AA; 7797 MW; 6853F12B33625B CRC64;

Query Match 0.7%; Score 8; DB 1; Length 76;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 12 GLLLLALL 19

RESULT 12
CD24_RAT
ID CD24_RAT STANDARD; PRT; 76 AA.
AC Q07490;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Signal transducer CD24 precursor (Heat stable antigen) (HSA)
DE (Nectadrin).
GN CD24A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
(1)
RN SEQUENCE FROM N.A.
RX STRAIN=Wistar; TISSUE=Embryonic brain;
RC MEDLINE=94122434; PubMed=8292828;
RA Shirasawa T., Akashi T., Sakamoto K., Takahashi H., Maruyama N.,
RA Hirokawa K.;
RT "Gene expression of CD24 core peptide molecule in developing brain
and developing non-neural tissues."
RL Dev. Dyn. 198:1-13(1993).
(2)
RN SEQUENCE FROM N.A.
RX STRAIN=Fischer;

```

```

RX MEDLINE=97157759; PubMed=9004038;
RA Magnaldo T.A., Barrandon Y.;
RT "CD24 (heat stable antigen, nectadrin), a novel keratinocyte
RT differentiation marker, is preferentially expressed in areas of the
RT hair follicle containing the colony-forming cells.";
RL J. Cell Sci. 109:3035-3045(1998).
CC
CC -!- FUNCTION: May have a pivotal role in cell differentiation. The
CC triggering mechanism of signal transduction may be due to the
CC interactions of differentiating cells with the matrix substrate
CC via the carbohydrate structure of the molecule. In this way, the
CC signal transducer can play very different roles in different cell
CC types as a direct consequence of its glycosylation.
CC
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC
CC -!- TISSUE SPECIFICITY: Expressed in the central nervous system, in
CC postmitotic cells of spinal cord, hindbrain, midbrain and
CC forebrain. Expressed in epithelium during the development of non-
CC neural tissues. Expressed in tooth development, specifically in
CC mesenchymal cells differentiating into odontoblast in dental
CC papilla, as well as in the developing eye and hair follicle.
CC
CC -!- DEVELOPMENTAL STAGE: Detected in primitive ectoderm, mesoderm and
CC ventral endoderm; down-regulated when organogenesis is completed.
CC
CC -!- PTM: Extensively O-glycosylated (by similarity). The carbohydrate
CC structure may be regulated in a tissue-specific and developmental
CC stage-specific manner.
CC
CC -!- SIMILARITY: TO OTHER MAMMALIAN SPECIES CD24.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z11663; CAA77731.1; -.
DR EMBL; U49062; AAA91470.1; -.
DR PIR; I53107; I53107.
KW Glycoprotein; GPI-anchor; Membrane; Signal; Differentiation;
KW Lipoprotein.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 56 SIGNAL TRANSDUCER CD24. (BY SIMILARITY).
FT PROPEP 57 76 REMOVED IN NATURE FORM (POTENTIAL).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 56 56 GPI-anchor amidated serine (Potential).
FT SEQUENCE 76 AA; 7862 MW; 4284670EC39D958 CRC64;
Query Match 0.7%; Score 8; DB 1; Length 76;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1152 GLLLLALL 1159
Db 12 GLLLLALL 19
RESULT 13
ID CD24_HUMAN STANDARD; PRT; 80 AA.
AC P25063; Q16257.
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Signal transducer CD24 precursor.
GN CD24 OR CD24A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91332458; PubMed=1831224;

```

```

RA Kay R., Rosten P.M., Humphries R.K.;
RT "CD24, a signal transducer modulating B cell activation responses, is
RT a very short peptide with a glycosyl phosphatidylinositol membrane
RT anchor.";
RL J. Immunol. 147:1412-1416(1991).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT VAL-57.
RX MEDLINE=93007871; PubMed=1327504;
RA Jackson D., Waibel R., Weber E., Bell J., Stabel R.A.;
RT "CD24, a signal-transducing molecule expressed on human B cells, is a
RT major surface antigen on small cell lung carcinomas.";
RL Cancer Res. 52:5264-5270(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feilsgold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Rosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.M.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinchman J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 1-76 FROM N.A.
RX MEDLINE=95048364; PubMed=7959762;
RA Hough M.R., Rosten P.M., Sexton T.L., Kay R., Humphries R.K.;
RT "Mapping of CD24 and homologous sequences to multiple chromosomal
RT loci.";
RL Genomics 22:154-161(1994).
CC -!- FUNCTION: Modulates B-cell activation responses. Signaling could
CC be triggered by the binding of a lectin-like ligand to the CD24
CC carbohydrates, and transduced by the release of second messengers
CC derived from the GPI-anchor. Promotes AG-dependent proliferation
CC of B-cells, and prevents their terminal differentiation into
CC antibody-forming cells.
CC
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC
CC -!- TISSUE SPECIFICITY: B-cells.
CC
CC -!- INDUCTION: Expression is lost when primary B-cells are induced to
CC differentiate in antibody-forming cells.
CC
CC -!- PTM: Extensively O-glycosylated.
CC
CC -!- SIMILARITY: TO OTHER MAMMALIAN SPECIES CD24.
CC
CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD24 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd24.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M58664; AAA35665.1; -.
DR EMBL; X69397; CAA49195.1; -.
DR EMBL; L33930; AAB58307.1; -.
DR EMBL; BC007674; AAB07674.1; -.
DR EMBL; S75311; AAD14170.1; ALT_INIT.
DR PIR; I56114; A48996.
DR Genew; HGNC:1645; CD24.

```

DR MM; 600074; --  
DR GO; GO:0006959; P:humoral immune response; TAS.  
KW Glycoprotein; GPI-anchor; Membrane; Signal; Polymorphism; Lipoprotein.  
FT SIGNAL 1 26 POTENTIAL.  
FT CHAIN 27 59 SIGNAL TRANSDUCER CD24.  
FT PROPEP 60 80 REMOVED IN MATURE FORM (POTENTIAL).  
FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT LIPID 59 59 GPI-anchor amidated glycine (Potential).  
FT VARIANT 57 57 A -> V (in dbSNP:8734).  
FT /FTID=VAR\_016156.  
FT CONFLICT 12 12 G -> W (IN REF. 4).  
FT CONFLICT 44 44 S -> T (IN REF. 4).  
SQ SEQUENCE 80 AA; 8083 MW; DB1988B6808F833P CRC64;  
  
Query Match 0.7%; Score 8; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 5.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1152 GLLLLALL 1159  
Db 12 GLLLLALL 19  
  
RESULT 14  
LST1 HUMAN STANDARD; PRT; 97 AA.  
AC Q00453; Q00454; Q13669; Q9UUR5; Q9UUR6; Q9UUR7; Q9UUR8;  
AC Q9UUS1;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DB Leukocyte specific transcript 1 protein (B144 protein).  
GN LST1 OR B144.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 3), TISSUE SPECIFICITY, AND INDUCTION.  
RC TISSUE=Blood;  
RX MEDLINE=96006565; PubMed=7590964;  
RA Holzinger I., de Baey A., Messer G., Kick G., Zwiersina H.,  
RA Weiss E.H.;  
RT "Cloning and genomic characterization of LST1: a new gene in the human  
TNF region.";  
RL Immunogenetics 42:315-322(1995).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 4), TISSUE SPECIFICITY, AND  
INDUCTION.  
RC TISSUE=Lymphoblast;  
RX MEDLINE=38035883; PubMed=9367684;  
RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,  
RA Weiss E.H.;  
RT "Complex expression pattern of the TNF region gene LST1 through  
differential regulation, initiation, and alternative splicing.";  
RL Genomics 45:591-600(1997).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORMS 5; 6; 7; 8 AND 9), FUNCTION, AND  
SUBCELLULAR LOCATION.  
RC TISSUE=Peripheral blood;  
RX MEDLINE=20171517; PubMed=10706767;  
RA Rollinger-Holzinger I., Bibl B., Pauly M., Grieser U., Hentges F.,  
RA Auer B., Fall G., Schratzberger P., Niederwieser D., Weiss E.H.,  
RA Zwierzina H.;  
RT "LST1: a gene with extensive alternative splicing and immunomodulatory  
function.";  
RL J. Immunol. 164:3169-3176(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,  
RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,  
RA Lasky S., Hood L.;

RT "Sequence of the human major histocompatibility complex class III  
region.";  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Hirakawa M., Yamaguchi H., Imai K., Shimada J.,  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP FUNCTION, AND SUBCELLULAR LOCATION.  
RX MEDLINE=21372017; PubMed=11478849;  
RA Raghunathan A., Sivakamasundari R., Wolenski J., Poddar R.,  
RA Weissman S.M.;  
RT "Functional analysis of B144/LST1: a gene in the tumor necrosis  
factor cluster that induces formation of long filopodia in eukaryotic  
cells.";  
RL Exp. Cell Res. 268:230-244(2001).  
CC -!- FUNCTION: Possible role in modulating immune responses. Isoforms 1  
and 2 have an inhibitory effect on lymphocyte proliferation.  
Induces morphological changes including production of filopodia  
and microspikes when overexpressed in a variety of cell types and  
may be involved in dendritic cell maturation.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Also detected in  
a perinuclear region corresponding to the localization of the  
Golgi apparatus and throughout the cytoplasm.  
CC -!- ALTERNATIVE PRODUCTS:  
Event-Alternative splicing; Named isoforms=9;  
Comment-Additional isoforms seem to exist;  
Name=1; Synonyms=LST1/A;  
IsoId=O00453-1; Sequence=Displayed;  
Name=2; Synonyms=LST1/C;  
IsoId=O00453-2; Sequence=VSP\_050578;  
Name=3; Synonyms=PLST1;  
IsoId=O00453-3; Sequence=VSP\_050579;  
Name=4; Synonyms=LST1/E;  
IsoId=O00453-4; Sequence=VSP\_050577;  
Name=5; Synonyms=LST1/K;  
IsoId=O00453-5; Sequence=VSP\_050584;  
Name=6; Synonyms=LST1/L;  
IsoId=O00453-6; Sequence=VSP\_050583;  
Name=7; Synonyms=LST1/J;  
IsoId=O00453-7; Sequence=VSP\_050586;  
Name=8; Synonyms=LST1/M;  
IsoId=O00453-8; Sequence=VSP\_050580;  
Name=9; Synonyms=LST1/N;  
IsoId=O00453-9; Sequence=VSP\_050582;  
CC -!- TISSUE SPECIFICITY: Expressed in lung, tonsil, thymus, placenta,  
kidney, fetal spleen, fetal liver and brain.  
CC -!- INDUCTION: By interferon gamma.  
CC -!- SIMILARITY: Belongs to the LST1 family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U00921; AAB57724.1; --  
DR EMBL; AF000424; AAB86998.1; --  
DR EMBL; AF000425; AAB86999.1; --  
DR EMBL; AF000426; AAB87000.1; --  
DR EMBL; Y18486; CAB59904.1; --  
DR EMBL; Y18487; CAB59905.1; --  
DR EMBL; Y18488; CAB60038.1; --  
DR EMBL; Y18489; CAB59906.1; --  
DR EMBL; Y18490; CAB59903.1; --  
DR EMBL; AF129756; AAD18090.1; --  
DR EMBL; AF000505; BAB61394.1; --  
DR GeneW; HGNC:14189; LST1.  
DR MIM; 109170; --  
DR GO; GO:0006955; P:immune response; NAS.  
DR GO; GO:0009653; P:immunogenesis; NAS.



```

OS Klebsiella aerogenes.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
CX NCBI_TaxID=28451;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W70;
RX MEDLINE=20566700; PubMed=11114933;
RA Kolko M.M., Kapetanovich L.A., Lawrence J.G.;
RT "Alternative pathways for siroheme synthesis in Klebsiella
RT aerogenes.";
RL J. Bacteriol. 183:328-335(2001).
CC -!- FUNCTION: Required for the cell division process (By similarity).
CC -!- SUBUNIT: May interact with ftsL (By similarity).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane
CC (potential). Colocalizes with ftsL to the division site (By
CC similarity).
CC -!- SIMILARITY: Belongs to the ftsB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; AF308468; AAG42461.1; --
DR HAMAP; MF_00599; --; 1.
DR InterPro; IPR007060; DivIC.
DR Pfam; PF04977; DivIC; 1.
KW Cell division; Transmembrane; Inner membrane; Coiled coil.
FT DOMAIN 1 3 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 4 21 POTENTIAL.
FT DOMAIN 22 105 PERIPLASMIC (POTENTIAL).
FT DOMAIN 31 74 COILED COIL (POTENTIAL).
SQ SEQUENCE 105 AA; 11951 MW; 530471363FD3D112A CRC64;

Query Match 0.7%; Score 8; DB 1; Length 105;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160
Db 6 LLLALLV 13

RESULT 17
IL22_HUMAN
ID IL22_HUMAN STANDARD; PRT; 179 AA.
AC Q9GZX6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Interleukin-22 precursor (IL-22) (IL-10-related T-cell-derived
DE inducible factor) (IL-TIP) (UNQ3099/PRO10096).
GN IL22 OR ILTIF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20420346; PubMed=10954742;
RA Dumontier L., Van Roost E., Colau D., Renaud J.-C.;
RT "Human interleukin-10-related T cell-derived inducible factor:
RT molecular cloning and functional characterization as an hepatocyte-
RT stimulating factor.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:10144-10149(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21069354; PubMed=11197690;
RA Dumontier L., Van Roost E., Colau D., Amey G., Michaux L.,

```

```

RA Renaud J.-C.;
RT "IL-TIP/IL-22: genomic organization and mapping of the human and mouse
RT genes.";
RL Genes Immun. 1:488-494(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20469498; PubMed=10875937;
RA Xie M.-H., Aggarwal S., Ho W.-H., Foster J., Zhang Z., Stinson J.,
RA Wood W.I., Goddard A.D., Gurney A.L.;
RT "Interleukin (IL)-22, a novel human cytokine that signals through the
RT interferon receptor-related proteins CRF2-4 and IL-22R.";
RL J. Biol. Chem. 275:31335-31339(2000).
RN [4]
RP SEQUENCE FROM N.A. AND VARIANT GLY-158
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=22887296; PubMed=12975309;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel S., Dowd P.,
RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heidens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R., Watanabe C., Wieand D., Woods K., Xie M.-H., Yansura D.,
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
RA Godowski P.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270(2003).
CC -!- FUNCTION: Cytokine that contributes to the inflammatory response
CC in vivo.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the IL-10 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; AJ277247; CAC06085.1; --
DR ENBL; AJ277248; CAC19409.1; --
DR ENBL; AF279437; AAG22064.1; --
DR ENBL; AF387519; AAK62468.1; --
DR ENBL; AF388890; AAK89249.1; --
DR Genew; HGNC:14900; IL22.
DR MIM; 605330; --
DR GO; GO:0005576; C:extracellular; IC.
DR GO; GO:0045518; F:interleukin-22 receptor binding; NAS.
DR GO; GO:0006953; P:acute-phase response; NAS.
DR GO; GO:0007267; P:cell-cell signaling; IC.
DR GO; GO:0006954; P:inflammatory response; NAS.
DR InterPro; IPR000098; Interleukin_10.
DR PROSITE; PS0520; INTERLEUKIN_10; 1.
KW Cytokine; Glycoprotein; Signal; Polymorphism.
FT SIGNAL 1 33 POTENTIAL.
FT CHAIN 34 179 INTERLEUKIN-22.
FT CARBOHYD 54 54 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 68 68 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 97 97 N-LINKED (GLCNAC... ) (POTENTIAL).
FT VARIANT 158 158 S -> G.
FT /FTID=VAR_013078.
SQ SEQUENCE 179 AA; 20011 MW; 3C35E64D60CF8767 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 179;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



QY 1153 LLLLLLV 1160  
DB 21 LLLLLLV 28

## RESULT 18

EMBL2\_CAVPO STANDARD; PRT; 234 AA.  
AC F35709;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Eosinophil granule major basic protein 2 precursor (MBP-2).  
GN MBP2.  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
OX NCBI\_TaxID=10141;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Eosinophil;  
RX MEDLINE=91224343; PubMed=2026266;  
RA Aoki I., Shindoh Y., Nishida T., Nakai S., Hong Y.-M., Mio M.,  
RA Saito T., Tasaka K.;  
RT "Comparison of the amino acid and nucleotide sequences between human  
RL FEBS Lett. 282:56-60(1991).  
CC -!- FUNCTION: MBP may play some important roles in the allergic  
CC reactions and inflammations, since MBP is capable of releasing  
CC histamine from mast cells and damaging the epithelial cells of  
CC bronchial tubes. Antiparasitic and antibiotic.  
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific  
CC granule (crystalloid core).  
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).  
CC  
CC EMBL; D00817; BAA00697.1; -;  
DR PIR; S15102; S15102.  
DR HSSP; P13727; 1H8U.  
DR InterPro; IPR002352; Emaior basic.  
DR InterPro; IPR001304; Lectin\_C.  
DR Pfam; PF00059; lectin\_C\_1.  
DR PRINTS; PR00770; EMAJORBASICP.  
DR SMART; SM00034; CLECT; 1.  
DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.  
DR PROSITE; PS00041; C-TYPE LECTIN 2; 1.  
DR Eosinophil; Signal; Immune response; Antibiotic; Lectin;  
KW Multigene family; Glycoprotein.  
FT SIGNAL 1 15 POTENTIAL.  
FT PROPEP 16 115 ACIDIC.  
FT CHAIN 116 234 EOSINOPHIL GRANULE MAJOR BASIC PROTEIN 2.  
FT DOMAIN 133 234 C-TYPE LECTIN (SHORT FORM).  
FT DISULFID 135 232 BY SIMILARITY.  
FT DISULFID 209 224 BY SIMILARITY.  
FT CARBOHYD 69 O-LINKED (GLYCOSAMINOGLYCAN) (BY  
FT SIMILARITY)  
SQ SEQUENCE 234 AA; 26140 MW; 7D926A942BF5116F CRC64;  
Query Match 0.7%; Score 8; DB 1; Length 234;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1153 LLLLLLV 1160  
DB 4 LLLLLLV 11

## RESULT 19

COMT RAT STANDARD; PRT; 264 AA.  
AC P22734;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Catechol O-methyltransferase, membrane-bound form (EC 2.1.1.6)  
DE {MB-COMT} [Contains: Catechol O-methyltransferase, soluble form  
DE {S-COMT}].  
GN COMT.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=94107221; PubMed=8280056;  
RA Tenhunen J., Ullmanen I.;  
RT "Production of rat soluble and membrane-bound catechol O-  
RL Biochem. J. 296:595-600(1993).  
CC [2]  
CC SEQUENCE OF 11-264 FROM N.A.  
CC MEDLINE=91033034; PubMed=2227437;  
CC Salminen M., Lundstroem K., Tilgmann C., Savolainen R., Kalkkinen N.,  
CC Ullmanen I.;  
CC "Molecular cloning and characterization of rat liver catechol-O-  
CC methyltransferase.";  
CC Gene 93:241-247(1990).  
RN [3]  
RP SEQUENCE OF 1-10 FROM N.A., AND CHARACTERIZATION OF THE TWO FORMS.  
RX MEDLINE=92111472; PubMed=1765063;  
RA Ullmanen I., Lundstroem K.;  
RT "Cell-free synthesis of rat and human catechol O-methyltransferase.  
RT Insertion of the membrane-bound form into microsomal membranes in  
RT vitro.";  
RL Eur. J. Biochem. 202:1013-1020(1991).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF SOLUBLE FORM.  
RX MEDLINE=94173341; PubMed=8127373;  
RA Vidgren J., Svensson L.A., Liljas A.;  
RT "Crystal structure of catechol O-methyltransferase.";  
RL Nature 368:354-358(1994).  
CC -!- FUNCTION: Catalyzes the O-methylation, and thereby the  
CC inactivation, of catecholamine neurotransmitters and catechol  
CC hormones. Also shortens the biological half-lives of certain  
CC neuroactive drugs, like L-DOPA, alpha-methyl DOPA and  
CC isoproterenol.  
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + catechol = S-  
CC adenosyl-L-homocysteine + guaiacol.  
CC -!- COFACTOR: Magnesium.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (isoform S-COMT). Type II  
CC membrane protein (isoform MB-COMT).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative initiation;  
CC Comment=2 isoforms, Membrane-bound/MB-COMT (shown here) and  
CC Soluble/S-COMT, are produced by alternative initiation;  
CC -!- PTM: The N-terminus is blocked.  
CC -!- SIMILARITY: TO OTHER MAMMALIAN CATECHOL-O-METHYLTRANSFERASE.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z12651; CAA78276.1; -;

```

DR EMBL: M60754; AAA40892.1; ALT INIT.
DR EMBL: M60753; AAA40891.1; ALT_INIT.
DR PIR: S22090; S22090.
DR FDB; IVID; 11-JUL-96.
DR InterPro; IPR002935; Methyltransf_3.
DR InterPro; IPR000051; SAM_bind.
DR Pfam; PF01596; Methyltransf_3; 1.
KW Transferase; Neurotransmitter degradation;
KW Catecholamine metabolism; Transmembrane; Signal-anchor; Magnesium;
KW Alternative initiation; 3D-structure.
FT CHAIN 1 264
FT CATECHOL-O-METHYLTRANSFERASE, ISOFORM
FT MEMBRANE-BOUND.
FT CHAIN 45 264
FT CATECHOL-O-METHYLTRANSFERASE, ISOFORM
FT SOLUBLE.
FT INIT MET 44 44
FT TRANSMEM 3 19
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT ACT_SITE 187 187
FT ACT_SITE 242 242
FT METAL 184 184
FT METAL 212 212
FT METAL 213 213
FT HELIX 48 59
FT TURN 62 63
FT TURN 65 78
FT TURN 79 79
FT STRAND 104 108
FT TURN 111 112
FT HELIX 114 120
FT TURN 121 122
FT TURN 125 126
FT STRAND 128 133
FT HELIX 136 149
FT TURN 150 151
FT TURN 152 154
FT STRAND 155 159
FT HELIX 162 165
FT HELIX 166 168
FT HELIX 169 173
FT STRAND 178 183
FT HELIX 187 189
FT HELIX 190 199
FT TURN 200 201
FT STRAND 203 211
FT HELIX 214 218
FT TURN 220 228
FT TURN 230 231
FT STRAND 232 240
FT TURN 242 243
FT STRAND 247 255
SQ SEQUENCE 264 AA; 29597 MW; F535DFP49C062854 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 264;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 9 GLLLLALL 16

RESULT 20
OPDE_PSEAE OPDE_PSEAE STANDARD; PRT; 402 AA.
AC Q01602;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription regulatory protein opde.
GN OPDE OR PA2219.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

```

```

OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1 / H103;
RX MEDLINE=93051259; PubMed=1427017;
RA Huang H., Siehn R.J., Bellido F., Rawling E., Hancock R.E.W.;
RT "Analysis of two gene regions involved in the expression of the
aeruginosa-specific, outer membrane porin OprD of Pseudomonas
aeruginosa."
RL FEMS Microbiol. Lett. 76:267-274(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=2043737; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
RA Hickey M.J., Brinkman P.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -!- FUNCTION: REGULATES THE EXPRESSION OF OPD WHICH ENCODES THE
IMIPENEM-SPECIFIC PORIN.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO B.SUBTILIS IPA-79D.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: Z14064; CAA78446.1; -.
DR EMBL: AB004648; AAG05607.1; -.
DR PIR: S23860; S23860.
DR InterPro; IPR007114; MPS.
DR InterPro; IPR005828; Sub_transporter.
DR Pfam; PF00083; sugar_trf_1.
DR PROSITE; PS00850; MFS; 1.
KW Transcription regulation; Transmembrane; Complete proteome.
FT TRANSMEM 22 42
FT TRANSMEM 60 80
FT TRANSMEM 86 106
FT TRANSMEM 108 128
FT TRANSMEM 147 167
FT TRANSMEM 170 190
FT TRANSMEM 220 240
FT TRANSMEM 256 276
FT TRANSMEM 296 316
FT TRANSMEM 318 338
FT TRANSMEM 348 368
FT TRANSMEM 375 395
SQ SEQUENCE 402 AA; 41592 MW; 0C5701C4AD2FDE16 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 402;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1147 GSTLGLL 1154
DB 360 GSTLGLL 367

RESULT 21
NSMA_HUMAN STANDARD; PRT; 423 AA.
ID NSMA_HUMAN
AC C60906; Q9BWR3;
DT 28-FEB-2003 (Rel. 41, Created)

```

28-FEB-2003 (Rel. 41, Last sequence update)  
 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Spingomyelin phosphodiesterase 2 (EC 3.1.4.12) (Neutral  
 DE spingomyelinase) (nsMase) (N-SMase) (Lyso-platelet activating factor-  
 DE phospholipase C) (Lyso-PAF-PLC).  
 GN SMPD2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND TISSUE SPECIFICITY.  
 RX MEDLINE=98188255; PubMed=9520418;  
 RA Tomiuk S., Hofmann K., Nix M., Zumbansen M., Stoffel W.;  
 RT "Cloned mammalian neutral sphingomyelinase: functions in sphingolipid  
 RT signaling?";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:3638-3643 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong J.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.X., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.B.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16399-16903 (2002).  
 RN [3]  
 RP CHARACTERIZATION.  
 RX MEDLINE=20076490; PubMed=1060884;  
 RA Sawai H., Domae N., Nagai N., Hannun Y.A.;  
 RT "Function of the cloned putative neutral sphingomyelinase as  
 RT lyso-platelet activating factor-phospholipase C.";  
 RL J. Biol. Chem. 274:38131-38139 (1999).  
 CC -!- FUNCTION: Converts sphingomyelin to ceramide. Hydrolyze 1-acyl-2-  
 CC lyso-sn-glycero-3-phosphocholine (lyso-PC) and 1-O-alkyl-2-lyso-  
 CC sn-glycero-3-phosphocholine (lyso-platelet activating factor). The  
 CC physiological substrate seems to be Lyso-PAF.  
 CC -!- CATALYTIC ACTIVITY: Sphingomyelin + H(2)O = N-acylsphingosine +  
 CC choline phosphate.  
 CC -!- COFACTOR: Magnesium.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- MISCELLANEOUS: This protein has an optimum pH of 6.5-7.5.  
 CC -!- SIMILARITY: Belongs to the neutral sphingomyelinase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AJ222801; CAA10995.1; -;  
 CC EMBL; BC000038; AAH00038.1; -;  
 CC Genew; HGNC:11121; SMPD2.  
 CC MIM; 603498; -;  
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.  
 CC GO; GO:0004767; P:sphingomyelin phosphodiesterase activity; TAS.

GO; GO:0006684; P:sphingomyelin metabolism; TAS.  
 DR InterPro; IPR005115; Exo endo phos.  
 DR Pfam; PF03372; Exo endo phos; 1  
 KW Hydrolase; Transmembrane; Magnesium  
 FT TRANSMEM 330 350 POTENTIAL.  
 FT METAL 49 49 MAGNESIUM (BY SIMILARITY).  
 FT SITE 180 180 SIMILARITY.  
 FT ACT SITE 272 272 L->P (IN RES. 2).  
 FT CONFLICT 3 3 GENERAL BASE (BY SIMILARITY).  
 SQ SEQUENCE 423 AA; 47592 MW; 05252A923E363171 CRC64;  
 Query Match 0.7%; Score 8; DB 1; Length 423;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1152 GLLLLALL 1159  
 Db 334 GLLLLALL 341  
 RESULT 22  
 ENGA MYCPE STANDARD; PRT; 444 AA.  
 ID Q8WH6; 2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE GTP-binding protein engA.  
 DE ENGA OR MYP2290.  
 GN Mycoplasma penetrans.  
 OS Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OC NCBI\_TaxID=28227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HF-2;  
 RX MEDLINE=22354719; PubMed=12466535;  
 RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenzi T., Furuya K.,  
 RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;  
 RT "The complete genomic sequence of Mycoplasma penetrans, an  
 RT intracellular bacterial pathogen in humans.";  
 RL Nucleic Acids Res. 30:5293-5300 (2002).  
 CC -!- FUNCTION: GTPase of unknown physiological role.  
 CC -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding  
 CC proteins. EngA subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AP004170; BAC44020.1; -;  
 CC HAMAP; MF 00195; -; 1.  
 DR InterPro; IPR005289; GTP-binding\_dom.  
 DR InterPro; IPR002917; MVR\_HSR1.  
 DR InterPro; IPR001806; Ras\_trnafrmg.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF01926; MVR\_HSR1; 1.  
 DR PRINTS; PR00449; RASTRNSFRMG.  
 DR TIGRFAMs; TIGR00650; MG442; 2.  
 DR TIGRFAMs; TIGR00231; small GTP; 2.  
 KW GTP-binding; Repeat; Complete proteome.  
 FT NP\_BIND 8 15 GTP 1 (POTENTIAL).  
 FT NP\_BIND 55 59 GTP 1 (POTENTIAL).  
 FT NP\_BIND 118 121 GTP 1 (POTENTIAL).  
 FT NP\_BIND 179 186 GTP 2 (POTENTIAL).  
 FT NP\_BIND 226 230 GTP 2 (POTENTIAL).  
 FT NP\_BIND 291 294 GTP 2 (POTENTIAL).  
 SQ SEQUENCE 444 AA; 51231 MW; A7B636F43D7CB1BE CRC64;

Query Match 0.7%; Score 8; DB 1; Length 444;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 IIVLDGS 172  
DB 258 IIVLDGS 265

## RESULT 23

HISX\_COREF  
ID HISX\_COREF STANDARD; PRT; 451 AA.  
AC Q8FNZO.  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Histidinol dehydrogenase (EC 1.1.1.23) (HDH).  
GN HISD OR CE2003.  
OS Corynebacterium efficiens.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
OX NCBI\_TaxID=152794;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;  
RX MEDLINE=22723752; PubMed=12840036;  
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,  
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,  
RA Gotojori T.;  
RT "Comparative complete genome sequence analysis of the amino acid  
RT replacements responsible for the thermostability of Corynebacterium  
RL Genome Res. 13:1572-1579 (2003).  
CC -!- FUNCTION: Catalyzes the sequential NAD-dependent oxidations of L-  
CC histidinol to L-histidinolaldehyde and then to L-histidine (by  
CC similarity).  
CC -!- CATALYTIC ACTIVITY: L-histidinol + 2 NAD(+) + H(2)O = L-histidine  
CC + 2 NADH.  
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
CC -!- PATHWAY: Histidine biosynthesis; ninth (last) step.  
CC -!- SIMILARITY: Belongs to the histidinol dehydrogenase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AP005220; BAC18813.1; -  
CC HAMAP; MF 01024; -; 1.  
DR InterPro; IPR001692; Histidinol dh.  
DR Pfam; PF00815; Histidinol dh. 1.  
DR PRINTS; PR00083; H0LDHGRGNASE  
DR PRODOM; PD002880; Histidinol\_dh; 1.  
DR TIGRFAMs; TIGR00069; hisd; 1.  
DR PROSITE; PS00611; HISOL DEHYDROGENASE; FALSE NEG.  
KW Histidine biosynthesis; Oxidoreductase; NAD; Metal-binding; Zinc;  
Complete proteome.  
FT ACT\_SITE 332 332 BY SIMILARITY.  
FT ACT\_SITE 333 333 BY SIMILARITY.  
FT METAL 263 263 ZINC (BY SIMILARITY).  
FT METAL 266 266 ZINC (BY SIMILARITY).  
FT METAL 366 366 ZINC (BY SIMILARITY).  
FT METAL 425 425 ZINC (BY SIMILARITY).  
SQ SEQUENCE 451 AA; 47770 MW; 5A2DSAA7B990ECCF CRC64;

Query Match 0.7%; Score 8; DB 1; Length 451;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 DEAAAKDI 341  
DB 402 DEAAAKDI 409

## RESULT 24

D4DR\_HUMAN  
ID D4DR\_HUMAN STANDARD; PRT; 467 AA.  
AC P21917;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE D(4) dopamine receptor (D(2C) dopamine receptor).  
GN DRD4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ALLELE D4.7).  
RX MEDLINE=92310588; PubMed=1319557;  
RA van Tol H.H., Wu C.M., Guan H.C., Ohara K., Bunzow J.R.,  
RA Civelli O., Kennedy J., Seeman P., Niznik H.B., Jovanovic V.,  
RT "Multiple dopamine D4 receptor variants in the human population.";  
RL Nature 358:149-152 (1992).  
RN [2]  
RP SEQUENCE FROM N.A. (ALLELE D4.2).  
RX MEDLINE=91204054; PubMed=1840645;  
RA van Tol H.H.M., Bunzow J.R., Guan H.-C., Sunahara R.K., Seeman P.,  
RA Niznik H.B., Civelli O.;  
RT "Cloning of the gene for a human dopamine D4 receptor with high  
RT affinity for the antipsychotic clozapine.";  
RL Nature 350:610-614 (1991).  
RN [3]  
RP 3D-STRUCTURE MODELING.  
RX MEDLINE=93038586; PubMed=1358063;  
RA Livingstone C.D., Strange P.G., Naylor L.H.;  
RT "Molecular modelling of D2-like dopamine receptors.";  
RL Biochem. J. 287:277-282 (1992).  
RN [4]  
RP VARIANT GLY-194.  
RX MEDLINE=95243275; PubMed=7726213;  
RA Seeman P., Ulpian C., Chouinard G., van Tol H.H.M., Dwosh H.,  
RA Lieberman P., Sminovitch K., Liu I.S.C., Waye J., Voruganti P.,  
RA Hudson C., Serjeant G.R., Masibay A.S., Seeman M.V.;  
RT "Dopamine D4 receptor variant, D4-glycine-194, in Africans, but not in  
RT Caucasians: no association with schizophrenia.";  
RL Am. J. Med. Genet. 54:384-390 (1994).  
CC -!- FUNCTION: This is one of the five types (D1 to D5) of receptors  
CC for dopamine. The activity of this receptor is mediated by G  
CC proteins which inhibit adenylyl cyclase.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- POLYMORPHISM: The number of repeats is highly polymorphic and  
CC varies among different alleles. The sequence shown is that of  
CC allele D4.7.  
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L12398; AAB59386.1; -  
CC EMBL; L12397; -; NOT\_ANNOTATED\_CDS.  
DR PIR; S15079; DYH04.  
DR Genew; HGNC:3025; DRD4.  
DR MIM; 128452; -  
DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
DR GO; GO:0004952; F:dopamine receptor activity; TAS.

```
DR GO; GO:0007166; P:cell surface receptor linked signal transdu. ...; TAS.
DR GO; GO:0007212; P:dopamine receptor signaling pathway; TAS.
DR GO; GO:0007268; P:synaptic transmission; TAS.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PS00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR F1.1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR F1.2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Polymorphism; Repeat.
FT DOMAIN 1 37 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 38 60 1 (POTENTIAL).
FT DOMAIN 61 70 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 71 93 2 (POTENTIAL).
FT DOMAIN 94 109 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 110 131 3 (POTENTIAL).
FT DOMAIN 132 151 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 152 175 4 (POTENTIAL).
FT DOMAIN 176 191 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 192 213 5 (POTENTIAL).
FT DOMAIN 214 394 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 395 417 6 (POTENTIAL).
FT DOMAIN 418 426 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 427 449 7 (POTENTIAL).
FT DOMAIN 450 467 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 468 486 7 X 16 AA APPROXIMATE TANDEM REPEATS.
FT DOMAIN 249 360 1.
FT REPEAT 249 264 2.
FT REPEAT 265 280 3.
FT REPEAT 281 296 4.
FT REPEAT 297 312 5.
FT REPEAT 313 328 6.
FT REPEAT 329 344 7.
FT REPEAT 345 360 8.
FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 108 185 BY SIMILARITY.
FT VARIANT 194 194 V -> G (in Afro-Caribbeans;
FT VARIANT 265 344 dSNP:1800443).
FT VARIANT 281 328 Missing (in allele D4.2).
FT VARIANT 329 329 Missing (in allele D4.4).
FT VARIANT 332 332 Missing (in allele D4.4).
FT VARIANT 332 332 Missing (in allele D4.4).
FT SEQUENCE 467 AA; 48360 MW; B6PF2E09269A02AF CRC64;
Query Match 0.7%; Score 8; DB 1; Length 467;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88
|||||
|||||
RESULT 25
SIL7_HUMAN STANDARD; PRT; 467 AA.
AC Q9Y286; Q9Y286; Q9Y286; Q9Y502;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sialic acid binding Ig-like lectin 7 precursor (Siglec-7) (Q9Y286)
DE Membrane protein (Adhesion inhibitory receptor molecule-1) (AIRM-1)
DE (p75) (D-siglec).
GN SIGLEC7 OR AIRM1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```

```
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Dendritic cell.
RX MEDLINE=20036547; PubMed=10567377;
RA Nicoll G., Ni J., Liu D., Klennerman P., Munday J., Dubock S.,
RA Mattei M.-G., Crocker P.R.;
RT "Identification and characterization of a novel siglec, siglec-7,
RT expressed by human natural killer cells and monocytes.";
RL J. Biol. Chem. 274:34089-34095(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), PHOSPHORYLATION, AND
RP INTERACTION WITH PTPN6.
RC TISSUE=Lymphoid;
RX MEDLINE=99429884; PubMed=10499918;
RA Falco M., Biassoni R., Bottino C., Vitale M., Sivori S.,
RA Augugliaro R., Moretta L., Moretta A.;
RT "Identification and molecular cloning of p75/AIRM1, a novel member of
RT the sialoadhesin family that functions as an inhibitory receptor in
RT human natural killer cells.";
RL J. Exp. Med. 190:793-802(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Dendritic cell;
RA Zhang W., Wan T., Cao X.;
RT "Characterization of a novel siglec from dendritic cells.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBSJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=20230158; PubMed=10764831;
RA Angata T., Varki A.;
RT "Siglec-7, a sialic acid-binding lectin of the immunoglobulin
RT superfamily.";
RL Glycobiology 10:431-438(2000).
RN [5]
RP FUNCTION.
RX MEDLINE=20079612; PubMed=10611343;
RA Vitale C., Romagnani C., Falco M., Ponte M., Vitale M., Moretta A.,
RA Baggiolupo A., Moretta L., Mingari M.C.;
RT "Engagement of p75/AIRM1 or CD33 inhibits the proliferation of normal
RT or leukemic myeloid cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:15091-15096(1999).
RN [6]
RP DISIALOGLANGLIOSIDE BINDING.
RX MEDLINE=21286126; PubMed=11389909;
RA Ito A., Handa K., Withers D.A., Satoh M., Hakomori S.;
RT "Binding specificity of siglec7 to disialogangliosides of renal cell
RT carcinoma: possible role of disialogangliosides in tumor
RT progression.";
RL FEBS Lett. 498:116-120(2001).
CC -!- FUNCTION: Putative adhesion molecule that mediates sialic-acid
CC dependent binding to cells. Preferentially binds to alpha2,3 and
CC alpha2,6-linked sialic acid. Also binds disialogangliosides
CC (disialogangliosyl globoside, disialyl lactotetraosylceramide and
CC disialyl GalNAc lactotetraosylceramide). The sialic acid
CC recognition site may be masked by cis interactions with sialic
CC acids on the same cell surface. In the immune response, may act as
CC an inhibitory receptor upon ligand induced tyrosine
CC phosphorylation by recruiting cytoplasmic phosphatase(s) via their
CC SH2 domain(s) that block signal transduction through
CC dephosphorylation of signaling molecules. Mediates inhibition of
CC natural killer cells cytotoxicity. May play a role in hemopoiesis.
CC Inhibits differentiation of CD34+ cell precursors towards
CC myelomonocytic cell lineage and proliferation of leukemic myeloid
CC cells (in vitro).
CC -!- SUBUNIT: Interacts with PTPN6/SHP-1 upon phosphorylation.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1; Synonyms=AIMM-1b;
CC IsoId=Q9Y286-1; Sequence=Displayed;
CC Name=2; Synonyms=AIMM-2;
CC IsoId=Q9Y286-2; Sequence=VSP_002555;
CC Name=3; Synonyms=AIMM-3;
CC IsoId=Q9Y286-3; Sequence=VSP_002556, VSP_002558;
CC [1]
```

Note=No experimental confirmation available;  
 Name=4;  
 isoId=Q9Y286-4; Sequence=VSP\_002557, VSP\_002558;  
 Note=No experimental confirmation available;  
 TISSUE SPECIFICITY: Predominantly expressed by resting and activated natural killer cells and at lower levels by granulocytes and monocytes. High expression found in placenta, liver, lung, spleen, and peripheral blood leukocytes.  
 !- DOMAIN: Contains 1 copy of a cytoplasmic motif that is referred to as the immunoreceptor tyrosine-based inhibitor motif (ITIM). This motif is involved in downmodulation of cellular responses. The phosphorylated ITIM motif binds to the SH2 domain of FIPN6/SHP-1.  
 !- PTM: Tyrosine phosphorylated.  
 !- SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLSC (sialic acid binding Ig-like lectin) family.  
 !- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 !- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.  
 !- CAUTION: Ref.3 sequence differs from that shown due to a frameshift in position 406.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; AF170485; AAF12759.1; -  
 DR EMBL; AJ007395; CAB46011.1; -  
 DR EMBL; AJ130710; CAB51126.1; -  
 DR EMBL; AJ130711; CAB51127.1; -  
 DR EMBL; AJ130712; CAB51128.1; -  
 DR EMBL; AJ130713; CAB51129.1; -  
 DR EMBL; AF178981; AAF43446.1; ALT\_FRAME.  
 DR EMBL; AF193441; AAF06790.1; -  
 DR Genew; HGNC:10876; SIGLEC7.  
 DR MTM; 604410; -  
 DR GO; GO:0005887; C-integral to plasma membrane; TAS.  
 DR GO; GO:0005930; P-lectin; TAS.  
 DR GO; GO:0004872; F-receptor activity; TAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00408; IGC2; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 2.  
 KW Cell adhesion; lectin; Transmembrane; Signal; Glycoprotein;  
 KW Immunoglobulin domain; Repeat; Phosphorylation; Alternative splicing.  
 FT SIGNAL 1 18  
 FT CHAIN 19 467  
 FT DOMAIN 19 353  
 FT TRANSMEM 354 376  
 FT DOMAIN 377 467  
 FT DOMAIN 39 122  
 FT DOMAIN 150 233  
 FT DOMAIN 240 336  
 FT SITE 435 440  
 FT DISULFID 46 106  
 FT DISULFID 168 217  
 FT DISULFID 276 320  
 FT CARBOHYD 105 105  
 FT CARBOHYD 142 142  
 FT CARBOHYD 165 165  
 FT CARBOHYD 229 229  
 FT CARBOHYD 235 235  
 FT CARBOHYD 242 242  
 FT CARBOHYD 260 260  
 FT CARBOHYD 334 334  
 FT CARBOHYD 334 334  
 FT VARSPLIC 145  
 GTSVPLHPTTSSVLTLPQHQHGTSLTCQVTLPGAV  
 TTNRTQLNVSYS -> D (in isoform 2).  
 /FTid=VSP\_002555.

FT VARSPLIC 145 145 A->B (in isoform 3).  
 FT /FTid=VSP\_002556.  
 FT A->G (in isoform 4).  
 FT /FTid=VSP\_002557.  
 FT Missing (in isoform 3 and isoform 4).  
 FT /FTid=VSP\_002558.  
 FT CONFLICT 42 42 V->A (IN REF. 4).  
 SQ SEQUENCE 467 AA; 51142 MW; 8AFB4462B001F52 CRC64;  
 Query Match 0.7%; Score 8; DB 1; Length 467;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 379 GVILGAVG 386  
 DB 350 GVILGAVG 357  
 RESULT 26  
 ID\_CPBK MOUSE STANDARD; PRT; 470 AA.  
 AC Q62397;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome P450 2B20 (EC 1.14.14.1) (CYP2B20) (P24) (Fragment).  
 GN CYP2B20.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;  
 CC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Liver;  
 RX MEDLINE=96428606; PubMed=8831708;  
 RA Daron M., Pautrel A., Marc N., Guillouzo A., Corcos L.;  
 RT "Isolation of a new mouse cDNA clone: hybrid form of cytochrome P450  
 2B10 and NADPH-cytochrome P450 oxidoreductase.";  
 RL Biochem. Biophys. Res. Commun. 226:900-905(1996).  
 CC !- FUNCTION: Cytochromes P450 are a group of heme-thiolate  
 CC monooxygenases. In liver microsomes, this enzyme is involved in an  
 CC NADPH-dependent electron transport pathway. It oxidizes a variety  
 CC of structurally unrelated compounds, including steroids, fatty  
 CC acids, and xenobiotics.  
 CC !- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
 CC oxidized flavoprotein + H(2)O.  
 CC !- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC !- TISSUE SPECIFICITY: EXPRESSED IN THE LIVER AS WELL AS IN KIDNEY,  
 CC LUNG AND INTESTINE.  
 CC !- SIMILARITY: Belongs to the cytochrome P450 family.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; X99715; CAA68051.1; ALT\_INIT.  
 DR HSSP; P00179; 1D76.  
 DR MGD; MGI:1202389; Cyp2b20.  
 DR InterPro; IPR001128; Cytochrome P450.  
 DR InterPro; IPR008068; EP450\_CYP2B.  
 DR Pfam; PF00067; P450; 1  
 DR PRINTS; PRO1685; EP450ICYP2B.  
 DR PROSITE; PS00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR OXIDOREDUCTASE; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum; Phosphorylation.  
 FT MOD\_RES 128 128 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
 FT METAL 436 436 IRON (HEME AXIAL LIGAND).  
 FT NON\_TER 470 470

```
SQ SEQUENCE 470 AA; 53357 MW; 8B9CF3E2EA622642 CRC64;
Query Match 0.7%; Score 8; DB 1; Length 470;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLLV 1160
DB 6 LLLLLLLV 13

RESULT 27
XVLC_PSEPU STANDARD; PRT; 487 AA.
AC P43503;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Benzaldehyde dehydrogenase [NAD+] (EC 1.2.1.28).
GN XVLC
OS Pseudomonas putida.
OG Plasmid TOL pMW0, and Plasmid TOL pMW3.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.
RC PLASMID-TOL pMW0; PubMed=7848591;
RX MEDLINE=95173094;
RA Inoue J., Shaw J.P., Reik M., Harayama S.;
RT "Overlapping substrate specificities of benzaldehyde dehydrogenase
(the xylC gene product) and 2-hydroxymuconic semialdehyde
dehydrogenase (the xylG gene product) encoded by TOL plasmid pMW0 of
Pseudomonas putida";
RT Pseudomonas putida";
RL J. Bacteriol. 177:1196-1201(1995).
RN [2]
RP SEQUENCE OF 1-53.
RC PLASMID-TOL pMW53;
RX MEDLINE=91113163; PubMed=1989592;
RA Chalmers R.M., Keen J.N., Fawson C.A.;
RT "Comparison of benzyl alcohol dehydrogenases and benzaldehyde
dehydrogenases from the benzyl alcohol and mandelate pathways in
Acinetobacter calcoaceticus and from the TOL-plasmid-encoded toluene
pathway in Pseudomonas putida. N-terminal amino acid sequences, amino
acid compositions and immunological cross-reactions.";
RL Biochem. J. 273:99-107(1991).
CC -1- CATALYTIC ACTIVITY: Benzaldehyde + NAD(+) + H(2)O = benzoate +
NADH.
CC -1- SUBUNIT: Homotetramer.
CC -1- MISCELLANEOUS: Optimal pH is 9.0.
CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U15151; AAA66218.1; -
CC EMBL; D63341; BAA09661.1; -
CC PIR; T47107; T47107.
CC R5SP; P51977; 1bXS.
CC InterPro; IPR002086; Aldehyde_dehydr.
CC Pfam; PF00171; aldedh; 1.
CC PROSITE; PS00070; ALDEHYDE_DEHYDR_GLU; 1.
CC PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD; Plasmid.
FT NP_BIND 232 237 NAD (ADP PART) (BY SIMILARITY).
FT ACT_SITE 254 254 BY SIMILARITY.
FT ACT_SITE 288 288 BY SIMILARITY.
SQ SEQUENCE 487 AA; 51897 MW; 093CB3E9487AF384 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 626 ALGNVAIL 633
DB 171 ALGNVAIL 178

RESULT 28
CPB1_RAT STANDARD; PRT; 491 AA.
AC P00176;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2B1 (EC 1.14.14.1) (CYP1B1) (P450-B) (P450-PB1 and
P450-PB2) (P450-LM2).
GN CYP2B1 OR CYP2B-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE OF 6-491 FROM N.A. (ISOZYMES PB1 AND PB2).
RX MEDLINE=8222224; PubMed=6953431;
RA Fujii-Kuriyama Y., Mizukami Y., Kawajiri K., Sogawa K., Muramatsu M.;
RT "Primary structure of a cytochrome P-450: coding nucleotide sequence
of phenobarbital-inducible cytochrome P-450 cDNA from rat liver.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2793-2797(1982).
RN [2]
RP REVISIONS TO 166; 292 AND 378 (ISOZYMES PB1 AND PB2).
RA Fujii-Kuriyama Y., Mizukami Y., Kawajiri K., Sogawa K., Muramatsu M.;
RL Proc. Natl. Acad. Sci. U.S.A. 79:5443-5443(1982).
RN [3]
RP SEQUENCE OF 1-22.
RX MEDLINE=7919411; PubMed=109438;
RA Botelho L.H., Ryan D.E., Levin W.;
RT "Amino acid compositions and partial amino acid sequences of three
highly purified forms of liver microsomal cytochrome P-450 from rats
treated with polychlorinated biphenyls, phenobarbital, or 3-
methylcholanthrene.";
RL J. Biol. Chem. 254:5635-5640(1979).
RN [4]
RP PHOSPHORYLATION.
RX MEDLINE=90059885; PubMed=2583091;
RA Fyverin W., Taniguchi H.;
RT "Phosphorylation of hepatic phenobarbital-inducible cytochrome
P-450";
RL EMBO J. 8:3003-3010(1989).
CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate
monooxygenases. In liver microsomes, this enzyme is involved in an
NADPH-dependent electron transport pathway. It oxidizes a variety
of structurally unrelated compounds, including steroids, fatty
acids, and xenobiotics.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- INDUCTION: By phenobarbital.
CC -1- PTM: Phosphorylation is accompanied by a decrease in enzyme
activity.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; J00719; AAA41024.1; -
```



```

DR EMBL; M3713; AAC42028.1; -.
DR PIR; A00176; O4RTPB.
DR PIR; A21162; O4RTP2.
DR HSSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01685; EP450ICYP2B.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Phosphorylation.
FT MOD_RES 128 128 PHOSPHORYLATION (BY PKA).
FT METAL 436 436 IRON (HEME AXIAL LIGAND).
FT VARIANT 303 303 S -> G (IN ISOZYME PB2).
FT VARIANT 321 322 AE -> IV (IN ISOZYME PB2).
FT VARIANT 337 337 L -> P (IN ISOZYME PB2).
FT VARIANT 339 339 T -> S (IN ISOZYME PB2).
FT VARIANT 344 344 S -> T (IN ISOZYME PB2).
SQ SEQUENCE 491 AA; 55933 MW; 74615501AD5497DD CRC64;

Query Match 0.7%; Score 8; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
Db 6 LLLALLLV 13

RESULT 29
CPB2_RAT STANDARD; PRT; 491 AA.
AC P04187; Q64582;
DT 20-MAR-1987 (Rel. 04, Created)
DT 15-JUL-1998 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2B2 (SC 1.14.14.1) (CYP1B2) (P450B) (P450 PB4).
GN CYP2B2 OR CYP2B-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
CX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83247397; PubMed=6306654;
RA Mizukami Y., Sogawa K., Suwa Y., Muramatsu M., Fujii-Kuriyama Y.;
RT "Gene structure of a phenobarbital-inducible cytochrome P-450 in rat liver."
RL Proc. Natl. Acad. Sci. U.S.A. 80:3958-3962(1983).
RN [2]
RP SEQUENCE
RX MEDLINE=86059379; PubMed=3877725;
RA Frey A.B., Waxman D.J., Kreibich G.;
RT "The structure of phenobarbital-inducible rat liver cytochrome P-450 isoenzyme PB-4. Production and characterization of site-specific antibodies."
RL J. Biol. Chem. 260:15253-15265(1985).
RN [3]
RP SEQUENCE OF 168-491 FROM N.A.
RX MEDLINE=84159487; PubMed=6689485;
RA Phillips I.R., Shephard E.A., Ashworth A., Rabin B.R.;
RT "Cloning and sequence analysis of a rat liver cDNA coding for a phenobarbital-inducible microheterogenous cytochrome P-450 variant: regulation of its messenger level by xenobiotics."
RL Gene 26:41-52(1983).
RN [4]
RP SEQUENCE OF 281-491 FROM N.A.
RX MEDLINE=83291091; PubMed=6688421;
RA Kumar A., Raphael C., Adesnik M.;
RT "Cloned cytochrome P-450 cDNA. Nucleotide sequence and homology to multiple phenobarbital-induced mRNA species."
RL J. Biol. Chem. 258:111280-11284(1983).

```

```

RN [5]
RP ERRATUM.
RA Kumar A., Raphael C., Adesnik M.;
RL J. Biol. Chem. 259:6039-6039(1984).
RN [6]
RP MEDLINE=86205943; PubMed=3458196;
RX Atchison M.L.; Adesnik M.;
RA "Gene conversion in a cytochrome P-450 gene family."
RL Proc. Natl. Acad. Sci. U.S.A. 83:2300-2304(1986).
RN [7]
RP SEQUENCE OF 385-491 FROM N.A.
RX MEDLINE=84153837; PubMed=6322758;
RA Affolter M., Anderson A.;
RT "Segmental homologies in the coding and 3' non-coding sequences of rat liver cytochrome P-450e and P-450b cDNAs and cytochrome P-450e-like genes."
RL Biochem. Biophys. Res. Commun. 118:655-662(1984).
RN [8]
RP SEQUENCE OF 1-20 FROM N.A.
RX MEDLINE=88273074; PubMed=2839467;
RA Hashimoto T., Matsumoto T., Nishizawa M., Kawabata S.,
RA Morohashi K., Handa S., Omura T.;
RT "A mutant rat strain deficient in induction of a phenobarbital-inducible form of cytochrome P-450 in liver microsomes."
RL J. Biochem. 103:487-492(1988).
RN [9]
RP PHOSPHORYLATION.
RX MEDLINE=90059885; PubMed=2583091;
RA Pyerin W., Taniguchi H.;
RT "Phosphorylation of hepatic phenobarbital-inducible cytochrome P-450."
RL EMBO J. 8:3003-3010(1989).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate monooxygenases. In liver microsomes, this enzyme is involved in an NADPH-dependent electron transport pathway. It oxidizes a variety of structurally unrelated compounds, including steroids, fatty acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH + oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: By phenobarbital.
CC -!- PTM: Phosphorylation is accompanied by a decrease in enzyme activity.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR EMBL; J00728; AAA41056.1; -.
DR EMBL; J00720; AAA41056.1; JOINED.
DR EMBL; J00721; AAA41056.1; JOINED.
DR EMBL; J00722; AAA41056.1; JOINED.
DR EMBL; J00723; AAA41056.1; JOINED.
DR EMBL; J00724; AAA41056.1; JOINED.
DR EMBL; J00725; AAA41056.1; JOINED.
DR EMBL; J00726; AAA41056.1; JOINED.
DR EMBL; K00996; AAA41029.1; -.
DR EMBL; K01626; AAA41037.1; -.
DR EMBL; K01721; AAA41026.1; -.
DR EMBL; D00250; BAA00181.1; -.
DR EMBL; M13234; AAA41057.1; -.
DR PIR; A21162; O4RTP2.
DR HSSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008068; EP450_CYP2B.

```

```
DR PFam; PF00067; p450; 1.
DR PRINTS; PR01685; EP450ICYP2B.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Phosphorylation
FT MOD_RES 128 128 PHOSPHORYLATION (BY P450).
FT METAL 128 128 IRON (HEME AXIAL LIGAND).
SQ SEQUENCE 491 AA; 55932 MW; 00C6S937FDD44BC CRC64;

Query Match 0.7%; Score 8; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160
DB 6 LLLLLLV 13

RESULT 30
CPBA_MOUSE
ID CPBA_MOUSE STANDARD; PRT; 500 AA.
AC P12791;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2B10 (EC 1.14.14.1) (CYP2B10) (Testosterone 16-alpha
DE hydroxylase) (P450-16-alpha) (Clone PF3/46).
GN CYP2B10 OR CYP2B-10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89118235; PubMed=3219345;
RA Noshiro M., Lakso M., Kawajiri K., Negishi M.;
RT "Rip locus: regulation of female-specific isozyme (I-P-450(16 alpha)
RT of testosterone 16 alpha-hydroxylase in mouse liver, chromosome
RT localization, and cloning of P-450 cDNA.";
RL Biochemistry 27:6434-6443(1988).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; M21856; AAA40425.1; -.
DR PIR; B31047; B31047.
DR HSSP; P00179; 1DT6.
MGD; MGI:88598; Cyp2b10.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008068; EP450_CYP2B.
PFam; PF00067; p450; 1.
```

```
DR PRINTS; PR01685; EP450ICYP2B.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; FALSE NEG.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Phosphorylation
FT MOD_RES 128 128 PHOSPHORYLATION (BY P450).
FT METAL 128 128 IRON (HEME AXIAL LIGAND).
SQ SEQUENCE 500 AA; 56743 MW; F660A00D9D0FBA94 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160
DB 6 LLLLLLV 13

RESULT 31
VE1_BP2
ID VE1_BP2 STANDARD; PRT; 604 AA.
AC P11298;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Replication protein E1.
GN E1
OS Bovine papillomavirus type 2.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10560;
RN [1]
RP SEQUENCE FROM N.A.
RA Groff D.E., Mitra R., Lancaster W.D.;
RL Submitted (MAY-1988) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF
CC VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
CC PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
CC CONTAINS BINDING SITES FOR BOTH PROTEINS.
CC -!- SUBCELLULAR LOCATION: Nuclear.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; M20219; AAA66833.1; -.
DR PIR; C31169; W1WLB2.
DR InterPro; IPR001177; Papillom_E1.
DR PFam; PF00519; E1; 1.
DR PFam; PF00524; E1_N; 1.
KW Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
KW Nuclear protein.
FT NP_BIND 432 439 ATP (POTENTIAL).
SQ SEQUENCE 604 AA; 68077 MW; D2D7036ADE88A9DD CRC64;

Query Match 0.7%; Score 8; DB 1; Length 604;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 927 LAAGSDSN 934
DB 336 LAAGSDSN 343

RESULT 32
VE1_BPVI
ID VE1_BPVI STANDARD; PRT; 605 AA.
AC P03116; Q9WMH1;
DT 21-JUL-1986 (Rel. 01, Created)
```







```

CC      gene model prediction.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: X73975; CAA52155.1; -
CC      EMBL: AE003491; AAF48242.1; ALT_SEQ.
CC      F01: S40311; S40311.
CC      HSP: P11215; IAB8.
CC      FlyBase: FBgn0004456; msw.
CC      GO: GO:0004895; F: cell adhesion receptor activity; IMP.
CC      GO: GO:0007476; P: wing morphogenesis; ISP.
CC      InterPro: IPR000413; Intergrin_alpha.
CC      Pfam: PF01839; FG-GAP 3.
CC      Pfam: PF00357; Integrin_A; 1.
CC      SMART: SM00191; Int_alpha; 5.
CC      PROSITE: PS00242; INTEGRIN_ALPHA; 1.
CC      KX: Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC      Signal; Repeat.
CC      SIGNAL: 1 30
CC      CHAIN: 31 1346
CC      CHAIN: 31 2960
CC      CHAIN: 2961 1146
CC      CHAIN: 1146
CC      DOMAIN: 31 1085
CC      TRANSMEM: 1086 1106
CC      DOMAIN: 1107 1146
CC      REPEAT: 49 116
CC      REPEAT: ? ?
CC      REPEAT: 199 263
CC      REPEAT: 264 317
CC      REPEAT: 318 374
CC      REPEAT: 377 434
CC      REPEAT: 443 497
CC      CARBOHYD: 68 68
CC      CARBOHYD: 86 86
CC      CARBOHYD: 147 147
CC      CARBOHYD: 470 470
CC      CARBOHYD: 511 511
CC      CARBOHYD: 657 657
CC      CARBOHYD: 680 680
CC      CARBOHYD: 711 711
CC      CARBOHYD: 718 718
CC      CARBOHYD: 761 761
CC      CARBOHYD: 928 928
CC      CARBOHYD: 1027 1027
CC      CONFLICT: 730 731
CC      CONFLICT: 815 815
CC      SEQUENCE: 1146 AA; 127973 MW; 89AB75BC8DE6854E CRC64;
CC      Query Match 0.7%; Score 8; DB 1; Length 1146;
CC      Best Local Similarity 100.0%; Pred. No. 44;
CC      Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC      QY 1151 GGLLHLL 1158
CC      DB 1094 GGLLHLL 1101
CC      RESULT 36
CC      ITAM_HUMAN
CC      ID ITAM_HUMAN STANDARD; PRT; 1152 AA.
CC      AC P11215;
CC      DT 01-JUL-1989 (Rel. 11, Created)
CC      DT 01-OCT-1996 (Rel. 34, Last sequence update)
CC      DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC      DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha

```

```

DE subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor Mol)
DE (Neutrophil adherence receptor).
GN ITGAM OR CR3A OR CD11B.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88315033; PubMed=2457584;
RA Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;
RT "The human leukocyte adhesion glycoprotein Mac-1 (complement receptor
RT type 3, CD11b) alpha subunit. Cloning, primary structure, and
RT relation to the integrins, von Willebrand factor and factor B."
RL J. Biol. Chem. 263:12403-12411(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88190151; PubMed=2833753;
RA Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;
RT "Molecular cloning of the alpha subunit of human and guinea pig
RT leukocyte adhesion glycoprotein Mol: chromosomal localization and
RT homology to the alpha subunits of integrins."
RL Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88257215; PubMed=2454931;
RA Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;
RT "Amino acid sequence of the alpha subunit of human leukocyte adhesion
RT receptor Mol (complement receptor type 3)."
RL J. Cell Biol. 106:2153-2158(1988).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93123748; PubMed=8419480;
RA Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;
RT "Structural analysis of the CD11b gene and phylogenetic analysis of
RT the alpha-integrin gene family demonstrate remarkable conservation of
RT genomic organization and suggest early diversification during
RT evolution."
RL J. Immunol. 150:480-490(1993).
RN [5]
RP SEQUENCE OF 9-1153 FROM N.A.
RX MEDLINE=89098893; PubMed=2563162;
RA Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L.,
RA Roth G.J.;
RT "cDNA sequence for the alpha M subunit of the human neutrophil
RT adherence receptor indicates homology to integrin alpha subunits."
RL Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).
RN [6]
RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92073318; PubMed=1683702;
RA Shelley C.S., Arnaout M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
RT developmentally regulated expression."
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7]
RP SEQUENCE OF 1-9 FROM N.A.
RX TISSUE=Blood;
RX MEDLINE=92144986; PubMed=1346576;
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT "Characterization of the myeloid-specific CD11b promoter."
RL Blood 79:865-870(1992).
RN [8]
RP SEQUENCE OF 17-31.
RX MEDLINE=87076671; PubMed=3539202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol:
RT conservation across species and homology to platelet IIb/IIIa."
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin

```

RT CR3 (CD11b/CD18).";  
RL Cell 80-631-638(1995).  
RN [10]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.  
RX MEDLINE=96363671; PubMed=8747460;  
RA Lee J.O., Bankston L.A., Arnaut M.A., Liddington R.C.;  
RT "Two conformations of the integrin  $\alpha$ -domain (I-domain): a pathway for  
RT activation?";  
RL Structure 3:1333-1340(1995).  
RN [11]  
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.  
RX MEDLINE=98362595; PubMed=9687375;  
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,  
RA Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrichson R.L.,  
RA Horton N.C., Kelley L.L., Midner A.M., Moon J.B., Mott J.E.,  
RA Muchler V.I., Tomich C.S., Watenpaugh K.D., Wiley V.H.;  
RT "Cation binding to the integrin CD11b I domain and activation model  
RT assessment.";  
RL Structure 6:923-935(1998).  
RN [12]  
RP 3D-STRUCTURE MODELING OF 17-616.  
RX MEDLINE=98226734; PubMed=9560195;  
RA Oxvig C., Springer T.A.;  
RT "Experimental support for a beta-propeller domain in integrin  $\alpha$ -  
RT subunits and a calcium binding site on its lower surface.";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).  
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS  
CC ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES  
CC AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.  
CC IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF  
CC THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D  
CC PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR  
CC FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES  
CC OF FIBRINOGEN GAMMA CHAIN.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M  
CC ASSOCIATES WITH BETA-2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND  
CC GRANULOCYTES.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -!- DATABASE: NAFB-PROM; NOTE-CD guide CD11b entry;  
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11b.htm".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J03925; AAA59544.1; -;  
DR EMBL; M18044; AAA59491.1; -;  
DR EMBL; J04145; AAA59303.1; -;  
DR EMBL; S52227; AAB24821.1; -;  
DR EMBL; S52152; AAB24821.1; JOINED.  
DR EMBL; S52153; AAB24821.1; JOINED.  
DR EMBL; S52154; AAB24821.1; JOINED.  
DR EMBL; S52155; AAB24821.1; JOINED.  
DR EMBL; S52157; AAB24821.1; JOINED.  
DR EMBL; S52159; AAB24821.1; JOINED.  
DR EMBL; S52161; AAB24821.1; JOINED.  
DR EMBL; S52164; AAB24821.1; JOINED.  
DR EMBL; S52165; AAB24821.1; JOINED.  
DR EMBL; S52167; AAB24821.1; JOINED.  
DR EMBL; S52169; AAB24821.1; JOINED.  
DR EMBL; S52170; AAB24821.1; JOINED.  
DR EMBL; S52173; AAB24821.1; JOINED.  
DR EMBL; S52174; AAB24821.1; JOINED.

DR EMBL; S52180; AAB24821.1; JOINED.  
DR EMBL; S52181; AAB24821.1; JOINED.  
DR EMBL; S52184; AAB24821.1; JOINED.  
DR EMBL; S52189; AAB24821.1; JOINED.  
DR EMBL; S52191; AAB24821.1; JOINED.  
DR EMBL; S52192; AAB24821.1; JOINED.  
DR EMBL; S52203; AAB24821.1; JOINED.  
DR EMBL; S52212; AAB24821.1; JOINED.  
DR EMBL; S52213; AAB24821.1; JOINED.  
DR EMBL; S52216; AAB24821.1; JOINED.  
DR EMBL; S52219; AAB24821.1; JOINED.  
DR EMBL; S52221; AAB24821.1; JOINED.  
DR EMBL; S52222; AAB24821.1; JOINED.  
DR EMBL; S52223; AAB24821.1; JOINED.  
DR EMBL; S52226; AAB24821.1; JOINED.  
DR EMBL; M76724; AAA58410.1; -;  
DR EMBL; M84477; AAA51960.1; -;  
DR FIR; A31108; RWHUB.  
DR PDB; 1A8X; 17-JUN-98.  
DR PDB; 1BHO; 18-NOV-98.  
DR PDB; 1BHO; 18-NOV-98.  
DR PDB; 1IDN; 25-NOV-98.  
DR PDB; 1IDO; 01-AUG-96.  
DR PDB; 1JL4; 11-JAN-97.  
DR PDB; 1MIU; 07-AUG-02.  
DR Genew; HGNC:6149; ITGAM.  
DR MIM; 120980; -;  
DR GO; GO:0008305; C: integrin complex; TAS.  
DR GO; GO:0004895; P: cell adhesion receptor activity; TAS.  
DR GO; GO:0007155; P: cell adhesion; TAS.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR00453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 4.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS00234; VWFA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Signal; 3D-structure; Repeat; Magnesium; Calcium.  
FT SIGNAL 1 16  
FT CHAIN 17 1152 INTEGRIN ALPHA-M.  
FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 1105 1128 POTENTIAL.  
FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).  
Query Match 0.7%; Score 8; DB 1; Length 1152;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1151 GGLLLAL 1158  
DB 1115 GGLLLAL 1122  
RESULT 37  
DD37\_HUMAN STANDARD; PRT; 1157 AA.  
AC Q81Y37; Q8BU17; Q9P211;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Probable ATP-dependent helicase DHX37 (DEAH-box protein 37).  
GN DHX37 OR DD37 OR KIAA1517.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.



RC TISSUE=Brain, and Pubmed=12477932;  
RX MEDLINE=22388257; Pubmed=12477932;  
RA Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek A.M., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny N.J., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton B., Kettman M., Mañan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J.J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE OF 178-1157 FROM N.A.  
RX TISSUE=Brain;  
RX MEDLINE=20277482; Pubmed=10819231;  
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XVII.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 7:143-150(2000).  
CC -!- SIMILARITY: Belongs to the DEAD box helicase family. DEAD  
CC subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; BC002575; AAH02575.1; ALT\_INIT.  
DR EMBL; BC037964; AAH37964.1; -.  
DR EMBL; AB040950; BAA96041.1; -.  
DR Genew; HGNC:17210; DHX37.  
DR InterPro; IPR001410; DEAD.  
DR InterPro; IPR001650; Helicase\_C.  
DR Pfam; PF00408; HA2; 1.  
DR Pfam; PF00271; helicase\_C; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR SMART; SM00490; HELIC; 1.  
DR PROSITE; PS00690; DEAD ATP HELICASE; FALSE\_NSG.  
DR KMW; KMW00000; Helicase; ATP-binding.  
DR NP BIND 275 282 ATP (POTENTIAL).  
DR SITE 372 375 DEAD BOX.  
DR SITE 372 375 S -> G (IN REF. 1; AAH02575).  
DR CONFLICT 869 869 A -> S (IN REF. 2).  
DR CONFLICT 898 898 R -> Q (IN REF. 1; AAH02575).  
DR CONFLICT 1081 1081 YLLAEYCEWLPQAMHPDIEKAMPPTVH -> CEFDQGGGV  
DR CONFLICT 1130 1157 GVDKMSLRQGLCALCTVSPGLAEGSGPTANGOLEAT (IN  
DR REF. 2).  
DR SQ SEQUENCE 1157 AA; 129544 MW; 4933217522183005 CRC64;  
Query Match 0.7%; Score 8; DB 1; Length 1157;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 380 VLLGVGCA 387  
DB 857 VLLGVGCA 864

RESULT 38  
ITAD\_HUMAN STANDARD; PRT; 1162 AA.  
AC Q13349; Q15575; Q15576;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).  
GN ITGAD.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Spleen;  
RX MEDLINE=96111956; Pubmed=8777714;  
RA Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,  
RA Staunton D.E., Gallatin W.M.;  
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-3.";  
RL J. Biol. Chem. 275:8959-8969(2000).  
RN [3]  
RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.  
RX MEDLINE=96252336; Pubmed=8666289;  
RA Wong D.A., Davis E.M., Lebeau M., Springer T.A.;  
RT "Cloning and chromosomal localization of a novel gene-encoding a human  
RT beta 2-integrin alpha subunit.";  
RL Gene 171:291-294(1996).  
RN [4]  
RP INTERACTION WITH VCAM1.  
RX MEDLINE=99059642; Pubmed=9841932;  
RA Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,  
RA Hoffman P.A., Staunton D.E., Bochner B.S.;  
RT "alpha2beta2 integrin is expressed on human eosinophils and functions  
RT as an alternative ligand for vascular cell adhesion molecule 1  
RT (VCAM-1).";  
RL J. Exp. Med. 188:2187-2191(1998).  
RN [5]  
RP INTERACTION WITH VCAM1.  
RX MEDLINE=99370002; Pubmed=10438935;  
RA Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,  
RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;  
RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a  
RT binding interface between I domain and VCAM-1.";  
RL J. Immunol. 163:1984-1990(1999).  
CC -!- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND  
CC VCAM1. MAY PLAY A ROLE IN THE ATHEROSCLEROTIC PROCESS SUCH AS  
CC CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-  
CC BORNE PATHOGENS, PARTICULATE MATTER, AND SENESCENT ERYTHROCYTES  
CC FROM THE BLOOD.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D  
CC ASSOCIATES WITH BETA-2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL  
CC LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON  
CC TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN  
CC ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSET) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE. family.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain domain.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration



entires requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

Qy 1151 GGLLAL 1158  
Db 1115 GGLLAL 1122

RESULT 40  
TSP1\_XENLA  
ID TSP1\_XENLA STANDARD; PRT; 1173 AA.  
AC P35418; 1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DB Thrombospondin 1 precursor.  
GN TBS1 OR TSP1.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodidae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN (1)  
RP SEQUENCE FROM N.A.  
RA Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;  
RL Submitted (XX-1993) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and  
cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,  
laminin, type V collagen and integrins alpha-V/beta-1, alpha-  
V/beta-3 and alpha-IIB/beta-3 (By similarity).  
CC -!- SUBUNIT: Homotrimer; disulfide-linked.  
CC -!- SIMILARITY: Belongs to the thrombospondin family.  
CC -!- SIMILARITY: Contains 1 WFPC domain.  
CC -!- SIMILARITY: Contains 3 EGF-like domains.  
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.  
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.  
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M81695; AA559180.1; -.  
EMBL; Y00093; CA68283.1; -.  
EMBL; M29155; -; NOT ANNOTATED CDS.  
EMBL; M29487; AA51620.1; ALT SEQ.  
EMBL; M29482; AA51620.1; JOINED.  
EMBL; M29483; AA51620.1; JOINED.  
EMBL; M29484; AA51620.1; JOINED.  
EMBL; M29485; AA51620.1; JOINED.  
EMBL; M29486; AA51620.1; JOINED.  
PIR; A36584; RWHULC.  
PDB; 1N3Y; 18-FEB-03.  
Genew; HGNC:6152; ITGAX.  
MIM; 151510; -.  
GO; GO:0008305; C:integrin complex; TAS.  
GO; GO:0004895; F:cell adhesion receptor activity; TAS.  
GO; GO:0004872; F:receptor activity; TAS.  
GO; GO:0007155; P:cell adhesion; TAS.  
GO; GO:0007397; P:histogenesis and organogenesis; TAS.  
InterPro; IPR000413; Integrin\_alpha.  
InterPro; IPR002035; VWF\_A.  
Pfam; PF01839; FG-GAP; 3.  
Pfam; PF00357; integrin\_A; 1.  
Pfam; PF00092; vwa; 1.  
PRINTS; PR01185; INTEGRINA.  
SMART; SM00453; VWFADOMAIN.  
SMART; SM00191; Int\_alpha; 5.  
SMART; SM00327; VWA; 1.  
PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
PROSITE; PS00234; VWFA; 1.  
Signal; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
Signal; Magnesium; Calcium; Repeat; 3D-structure.  
SIGNAL 1 19  
CHAIN 20 1163 INTEGRIN\_ALPHA-X.  
DOMAIN 20 1107 EXTRACELLULAR (POTENTIAL).  
TRANSMEM 1108 1128 POTENTIAL.  
DOMAIN 1129 1163 CYTOPLASMIC (POTENTIAL).  
REPEAT 34 87 FG-GAP 1.  
REPEAT ? ? FG-GAP 2.  
DOMAIN 165 351 VWFA.  
REPEAT ? ? FG-GAP 3.  
REPEAT 402 453 FG-GAP 4.  
REPEAT 455 517 FG-GAP 5.  
REPEAT 518 576 FG-GAP 6.  
REPEAT 581 633 FG-GAP 7.  
CA\_BIND 466 474 POTENTIAL.  
CA\_BIND 530 538 POTENTIAL.  
CA\_BIND 593 601 POTENTIAL.  
SITE 1131 1135 GFFER MOTIF.  
DISULFID 68 76 BY SIMILARITY.  
DISULFID 108 126 BY SIMILARITY.  
DISULFID 655 712 BY SIMILARITY.  
DISULFID 771 777 BY SIMILARITY.  
DISULFID 848 863 BY SIMILARITY.  
DISULFID 998 1022 BY SIMILARITY.  
DISULFID 1027 1032 BY SIMILARITY.  
CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).  
CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).  
CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).  
CARBOHYD 697 697 N-LINKED (GLCNAC. .) (POTENTIAL).  
CARBOHYD 735 735 N-LINKED (GLCNAC. .) (POTENTIAL).  
CARBOHYD 899 899 N-LINKED (GLCNAC. .) (POTENTIAL).  
CARBOHYD 939 939 N-LINKED (GLCNAC. .) (POTENTIAL).  
CARBOHYD 1050 1050 N-LINKED (GLCNAC. .) (POTENTIAL).  
CONFLICT 490 490 G -> A (IN REF. 2).  
CONFLICT 756 756 L -> D (IN REF. 2).  
SEQUENCE 1163 AA; 127885 MW; 6C4E19CC3F62A473 CRC64;  
Query Match 0.7%; Score 8; DB 1; Length 1163;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|                       |   |   |               |           |                                     |
|-----------------------|---|---|---------------|-----------|-------------------------------------|
| KW                    | EGF-like domain; Signal.                      |   |               |           |                                     |
| FT                    | SIGNAL  | 1   | 22            |           | POTENTIAL.                          |
| FT                    | CHAIN   | 23  | 1173          |           | THROMBOSPONDIN 1.                   |
| FT                    | DOMAIN  | 23  | 235           |           | HEPARIN-BINDING (POTENTIAL).        |
| FT                    | DOMAIN  | 23  | 224           |           | TSP N-TERMINAL.                     |
| FT                    | DOMAIN  | 319   | 376           |           | WFCC.                               |
| FT                    | DOMAIN  | 382   | 432           |           | TSP TYPE-1 1.                       |
| FT                    | DOMAIN  | 438   | 493           |           | TSP TYPE-1 2.                       |
| FT                    | DOMAIN  | 495   | 550           |           | TSP TYPE-1 3.                       |
| FT                    | DOMAIN  | 550   | 590           |           | EGF-LIKE 1.                         |
| FT                    | DOMAIN  | 591   | 648           |           | EGF-LIKE 2.                         |
| FT                    | DOMAIN  | 649   | 693           |           | EGF-LIKE 3.                         |
| FT                    | DOMAIN  | 726   | 761           |           | TSP TYPE-3 1.                       |
| FT                    | DOMAIN  | 762   | 784           |           | TSP TYPE-3 2.                       |
| FT                    | DOMAIN  | 785   | 820           |           | TSP TYPE-3 3.                       |
| FT                    | DOMAIN  | 821   | 843           |           | TSP TYPE-3 4.                       |
| FT                    | DOMAIN  | 844   | 881           |           | TSP TYPE-3 5.                       |
| FT                    | DOMAIN  | 882   | 917           |           | TSP TYPE-3 6.                       |
| FT                    | DOMAIN  | 918   | 953           |           | TSP TYPE-3 7.                       |
| FT                    | DOMAIN  | 954   | 1173          |           | C-TERMINAL.                         |
| FT                    | SITE  | 929   | 931           |           | CELL ATTACHMENT SITE (POTENTIAL).   |
| FT                    | DISULFID                                      | 394   | 426           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 398   | 431           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 409   | 416           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 450   | 487           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 454   | 492           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 465   | 477           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 507   | 544           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 511   | 549           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 522   | 534           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 554   | 565           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 559   | 575           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 578   | 589           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 595   | 611           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 602   | 620           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 623   | 647           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 653   | 666           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 660   | 679           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 681   | 692           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 708   | 716           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 721   | 741           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 757   | 777           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 780   | 800           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 816   | 836           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 839   | 859           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 877   | 897           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 913   | 933           |           | BY SIMILARITY.                      |
| FT                    | DISULFID                                      | 949   | 1170          |           | BY SIMILARITY.                      |
| FT                    | CARBOHYD                                      | 155   | 155           |           | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT                    | CARBOHYD                                      | 158   | 158           |           | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT                    | CARBOHYD                                      | 250   | 250           |           | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT                    | CARBOHYD                                      | 363   | 363           |           | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT                    | CARBOHYD                                      | 705   | 705           |           | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT                    | CARBOHYD                                      | 711   | 711           |           | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT                    | CARBOHYD                                      | 1070  | 1070          |           | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| SQ                    | SEQUENCE                                      | 1173 AA; 130019 MW; A9F036D6516C0F24 CRC64; |               |           |                                     |
| Query Match           |   |   |               |           |                                     |
| Best Local Similarity |   | 0.7%; Score 8; DB 1; Length 1173;           |               |           |                                     |
| Matches               |   | 8; Conservative                             | 0; Mismatches | 0; Indels | 0; Gaps                             |
| OY                    | 486 SVIDGDG                                   | 493   |               |           |                                     |
| DB                    | 801 SVIDGDG                                   | 808   |               |           |                                     |
| RESULT 41             |   |   |               |           |                                     |
| ID                    | ITAE_HUMAN                                    | STANDARD;                                   | PRT;          | 1179 AA.  |                                     |
| AC                    | F38570; Q9NZU9;                               |   |               |           |                                     |
| DT                    | 01-OCT-1994 (Rel. 30, Created)                |   |               |           |                                     |
| DT                    | 30-MAY-2000 (Rel. 39, Last sequence update)   |   |               |           |                                     |
| DT                    | 28-FEB-2003 (Rel. 41, Last annotation update) |   |               |           |                                     |

|    |  |  |
|----|--|--|
| DE | Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1           |  |
| DE | antigen) (CD103 antigen) (Integrin alpha-IEL).                             |  |
| GN | ITGAE.   |  |
| OS | Homo sapiens (Human).  |  |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;          |  |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.                  |  |
| OX | NCBI_TaxID=9606;   |  |
| RN | [1]  |  |
| RP | SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.                     |  |
| RC | TISSUE=Leukemia, and Lymphocytes;  |  |
| RX | MEDLINE=94164962; PubMed=8119947;  |  |
| RA | Shaw S.K., Cepsek K.L., Murphy E.A., Russell G.J., Brenner M.B.,           |  |
| RA | Parker C.M.;   |  |
| RT | "Molecular cloning of the human mucosal lymphocyte integrin alpha E        |  |
| RT | subunit. Unusual structure and restricted RNA distribution.";              |  |
| RL | J. Biol. Chem. 269:6016-6025 (1994).                                       |  |
| RN | [2]  |  |
| RP | REVISIONS TO 88-114.   |  |
| RA | Parker C.M.;   |  |
| RL | Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.                    |  |
| RN | [3]  |  |
| RP | SEQUENCE OF 53-1179 FROM N.A.  |  |
| RC | TISSUE=Fetal kidney;   |  |
| RX | MEDLINE=20138496; PubMed=10673275;   |  |
| RA | Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G.,       |  |
| RA | Shettersuk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A.,         |  |
| RA | Green E.D.;  |  |
| RT | "The genomic region encompassing the nephropathic cystinosis gene          |  |
| RT | (CTNS): complete sequencing of a 200-kb segment and discovery of a         |  |
| RT | novel gene within the common cystinosis-causing deletion.";                |  |
| RL | Genome Res. 10:165-173 (2000).   |  |
| RN | [4]  |  |
| RP | MUTAGENESIS OF ASP-109 AND PHE-316.  |  |
| RX | MEDLINE=20400502; PubMed=10837471;   |  |
| RA | Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y.,       |  |
| RA | Brenner M.B.;  |  |
| RT | "The role of alpha and beta chains in ligand recognition by beta 7         |  |
| RT | integrins.";   |  |
| RL | J. Biol. Chem. 275:25652-25664 (2000).                                     |  |
| CC | -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT     |  |
| CC | MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL          |  |
| CC | CELL MONOLAYERS.   |  |
| CC | -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA         |  |
| CC | SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A             |  |
| CC | DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.                            |  |
| CC | -!- SUBCELLULAR LOCATION: Type I membrane protein.                         |  |
| CC | -!- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN     |  |
| CC | AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL          |  |
| CC | EPITHELIAL CELLS.  |  |
| CC | -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS     |  |
| CC | WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.                           |  |
| CC | -!- SIMILARITY: Belongs to the integrin alpha chain family.                |  |
| CC | -!- SIMILARITY: Contains 1 VWFA domain.                                    |  |
| CC | -!- SIMILARITY: Contains 7 FG-GAP repeats.                                 |  |
| CC | -!- DATABASE: NAME=PROW; NOTE=CD guide CD103 entry;                        |  |
| CC | WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd103.htm".                       |  |
| CC | -----  |  |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration |  |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation -    |  |
| CC | the European Bioinformatics Institute. There are no restrictions on its    |  |
| CC | use by non-profit institutions as long as its content is in no way         |  |
| CC | modified and this statement is not removed. Usage by and for commercial    |  |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ |  |
| CC | or send an email to license@isb-sib.ch)                                    |  |
| CC | -----  |  |
| DR | EMBL; L25951; AAB59359.2; ..   |  |
| DR | EMBL; AF168787; AAF43107.1; ..   |  |
| DR | PIR; A53213; A53213.   |  |
| DR | HSSP; P11215; 1A8X.  |  |
| DR | Genew; HGNC:6147; ITGAE.   |  |
| DR | MIM; 604682; ..  |  |
| DR | GO; GO:0008305; C:integrin complex; TAS.                                   |  |
| DR | GO; GO:0004895; F:cell adhesion receptor activity; TAS.                    |  |



RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Fulton L.;  
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP REVISIONS, AND ALTERNATIVE SPLICING.  
RA Waterston R.;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
RX MEDLINE=21585681; PubMed=11729150;  
RA Xu L., Strone S.;  
RL "Depletion of a novel SET-domain protein enhances the sterility of  
mes-3 and mes-4 mutants of *Caenorhabditis elegans*.";  
RT Genetics 159:1019-1029(2001).  
CC -1- FUNCTION: Probably involved in chromatin modification and/or  
remodelling in meiotic germ cells. May act redundantly with mes-3  
and mes-4 proteins.  
CC -1- SUBCELLULAR LOCATION: Nuclear. Localized in mitotic and mid-late-  
stage meiotic nuclei but is undetectable in early pachytene  
nuclei.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=a; Synonyms=L;  
CC IsoId=Q18221-1; Sequence=Displayed;  
CC Name=b; Synonyms=S;  
CC IsoId=Q18221-2; Sequence=VSP\_007217, VSP\_007218;  
CC -1- TISSUE SPECIFICITY: Expressed in all cells of embryo. In L1 larva,  
it is predominantly expressed in Z2 and Z3 primordial germ cells.  
CC In adult is predominantly expressed in the germline.  
CC -1- DEVELOPMENTAL STAGE: Expressed throughout embryogenesis.  
CC -1- SIMILARITY: Contains 1 post-SET domain.  
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.  
CC -1- SIMILARITY: Contains 1 SET domain.  
CC -1- CAUTION: Ref.1 (AAK67215) sequence differs from that shown due to  
erroneous gene model prediction.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U13875; AAA21163.1; -;  
CC EMBL; U13875; AAK67214.1; -;  
CC PIR; A88445; A88445  
CC Wormpep; C2656.9a; CE27735.  
CC Wormpep; C2656.9b; CEU1158.  
CC InterPro; IPR003616; PostSET.  
CC InterPro; IPR000504; RNA\_rec\_mot.  
CC InterPro; IPR001214; SET.  
CC Pfam; PF00856; SET; 1.  
CC SMART; SM00508; PostSET; 1.  
CC SMART; SM00360; RRM; 1.  
CC SMART; SM00317; SET; 1.  
CC PROSITE; PS50869; POST SET; 1.  
CC PROSITE; PS50102; RRM; FALSE NEG.  
CC PROSITE; PS00030; RRM\_RNP\_1; FALSE NEG.  
CC PROSITE; PS50280; SET; 1.  
KW RNA-binding; Nuclear protein; Alternative splicing.  
FT DOMAIN 128 199 RNA-BINDING (RRM) (POTENTIAL).  
FT DOMAIN 1366 1489 SET.  
FT DOMAIN 1491 1507 POST-SET.  
FT DOMAIN 296 354 PRO-RICH.  
FT DOMAIN 554 664 PRO-RICH.  
FT DOMAIN 870 1011 SER-RICH.  
FT VARSPLIC 1 768 Missing (in isoform b).  
FT /FTID=VSP\_007217.  
FT MDLSRKVAEDINQIMRCQFAALDKHLKAIADKKKK  
FT EREKRAQEAKEPSNLIADMM -> MYNNSAPYLNHSSLN

FT TVRKQVTVRRVLPSPPPPPPPPSLYPCSVKVPYIPQR  
FT VYRSINS (in isoform b).  
FT /FTID=VSP\_007218.  
SQ SEQUENCE 1507 AA; 171681 MW; E7D9689DA720C34A CRC64;  
Query Match 0.7%; Score 8; DB 1; Length 1507;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 430 TSVVSSRQ 437  
Db 1002 TSVVSSRQ 1009  
RESULT 44  
RPAL YEAST STANDARD; PRT: 1664 AA.  
AC P10964; Q99330;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE DNA-directed RNA polymerase I 190 kDa polypeptide (EC 2.7.7.6) (A190).  
GN RPAL OR RPA190 OR RRN1 OR YOR341W OR O6276.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88139333; PubMed=2830265;  
RA Memet S., Gouy M., Marck C., Sentenac A., Buhler J.-M.;  
RT "RPA190, the gene coding for the largest subunit of yeast RNA  
polymerase A";  
RL J. Biol. Chem. 263:2830-2839(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / FY73;  
RX MEDLINE=97103776; PubMed=8948102;  
RA Purnelle B., Goffeau A.;  
RT "Nucleotide sequence analysis of a 40 kb segment on the right arm of  
yeast chromosome XV reveals 18 open reading frames including a new  
pyruvate kinase and three homologues to chromosome I genes";  
RL Yeast 12:1475-1481(1996).  
CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription  
of DNA into RNA using the four ribonucleoside triphosphates as  
substrates. RNA polymerase I is essentially used to transcribe  
ribosomal DNA units.  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
{RNA} (N)  
CC -1- SUBUNIT: RNA polymerase I consists of 14 different subunits. This  
subunit is the largest component of RNA polymerase I.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases were  
found in eukaryotic nuclei: polymerase I for the ribosomal RNA  
precursor, polymerase II for the mRNA precursor, and polymerase  
III for 5S and tRNA genes.  
CC -1- SIMILARITY: Belongs to the RNA polymerase beta' chain family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; J03530; AAA34890.1; -;  
CC EMBL; X95720; CAA65029.1; -;  
CC EMBL; Z75249; CAA99665.1; -;  
CC PIR; S67250; S67250.  
CC GeneOnline; 143929; ...  
CC  
CC SGD; S0005868; RPA190.  
CC InterPro; IPR000722; RNA\_pol\_A.

```
DR InterPro; IPR007080; RNA_pol_Rpb1_1.
DR InterPro; IPR007086; RNA_pol_Rpb1_3.
DR InterPro; IPR007083; RNA_pol_Rpb1_4.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR InterPro; IPR006592; RNA_pol_A_N.
DR Pfam; PF04997; RNA_pol_Rpb1_1; 1.
DR Pfam; PF00623; RNA_pol_Rpb1_2; 1.
DR Pfam; PF04983; RNA_pol_Rpb1_3; 1.
DR Pfam; PF05000; RNA_pol_Rpb1_4; 1.
DR Pfam; PF04998; RNA_pol_Rpb1_5; 1.
DR SMART; SMO0663; RPOLA_N; 1.
KW Transferase; DNA-directed RNA polymerase; Transcription; Zinc;
XZ Zinc-finger; Nuclear protein.
FT ZN_FING 62 78 C2H2-TYPE (POTENTIAL).
FT CONFLICT 158 158 N->T (IN REF. 1).
SQ SEQUENCE 1664 AA; 186431 MW; DF6SAVAA459DSE6D CRC64;

Query Match 0.7%; Score 8; DB 1; Length 1664;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 481 GSEITSVD 488
DB 8 GSEITSVD 15

RESULT 45
RRPL_BUNYV
ID_RRPL_BUNYV STANDARD; PRT; 2238 AA.
AC P20470;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE RNA-directed RNA polymerase (EC 2.7.7.48) (L protein).
GN L.
OS Bunyamwera virus.
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Orthobunyavirus.
OX NCBI_TaxID=35304;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90085791; PubMed=2596023;
RA Elliott R.M.;
RT "Nucleotide sequence analysis of the large (L) genomic RNA segment of
Bunyamwera virus, the prototype of the family Bunyaviridae.";
RL Virology 173:426-436(1989).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
{RNA}(N).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X14383; CAA32553.1; -.
CC PIR; A33744; RRVUBY.
DR InterPro; IPR007322; Bunya_RdRp.
DR InterPro; IPR007099; RNA_pol_NSVir.
DR Pfam; PF04196; Bunya_RdRp; 1.
KW Transferase; Nucleotidyltransferase; RNA-directed RNA polymerase.
SQ SEQUENCE 2238 AA; 258668 MW; 1ED00AB156BAC8DA CRC64;

Query Match 0.7%; Score 8; DB 1; Length 2238;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 FLINILKK 191
DB 1333 FLINILKK 1340
```

```
RESULT 46
LSTL_MOUSE
ID LSTL_MOUSE STANDARD; PRT; 95 AA.
AC O08843; O08844; O9ZIH3; O9ZIR0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Leukocyte specific transcript 1 protein (B144 protein).
GN LSTL OR B144.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE-B-cell;
RX MEDLINE=88031493; PubMed=3117682;
RA Tsuge I., Shen F.-W., Steinmetz M., Boyse R.A.;
RT "A gene in the H-2S.H-2D interval of the major histocompatibility
complex which is transcribed in B cells and macrophages.";
RL Immunogenetics 26:378-380(1987).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN-BALB/c; TISSUE-Macrophage;
RX MEDLINE=98035883; PubMed=9367684;
RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
RA Weiss E.H.;
RT "Complex expression pattern of the TNF region gene LSTl through
differential regulation, initiation, and alternative splicing.";
RL Genomics 45:591-600(1997).
RN [3]
RP SEQUENCE FROM N.A.
RA Rowen L., Qin S., Madan A., Abbasi N., James R., Dickhoff R.,
RA Shaffer T., Ratcliffe A., Loretz C., Lasky S., Hood L.;
RT "Sequence of the mouse major histocompatibility class III region.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE=21372017; PubMed=11478849;
RA Raghunathan A., Sivakamasundari R., Wolenski J., Poddar R.,
RA Weissman S.M.;
RT "Functional analysis of B144/LSTl: a gene in the tumor necrosis
factor cluster that induces formation of long filopodia in eukaryotic
cells.";
RL Exp. Cell Res. 268:230-244(2001).
CC -!- FUNCTION: Possible role in modulating immune responses. Has an
inhibitory effect on lymphocyte proliferation. Induces
morphological changes including production of filopodia and
microspikes when overexpressed in a variety of cell types and may
be involved in dendritic cell maturation.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Also detected in
a perinuclear region corresponding to the localization of the
Golgi apparatus and throughout the cytoplasm (By
similarity).
CC -!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1; Synonyms=ml7r;
IsoId=O08843-1; Sequence=Displayed;
Name=2; Synonyms=ml21r;
IsoId=O08843-2; Sequence=VSP_050588;
CC -!- TISSUE SPECIFICITY: Expressed in spleen and at lower levels in
thymus and liver.
CC -!- SIMILARITY: Belongs to the LSTl family.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
frameshift in position 58.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to erroneous
gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```



CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
DR EMBL; M18187; AAA37273.1; ALT\_FRAME.  
DR EMBL; AF000427; AAB87001.1; -.  
DR EMBL; AF000428; AAB87002.1; -.  
DR EMBL; AF109719; AAC82482.1; ALT\_SEQ.  
DR PIR; 149515; I49515.  
DR MED; MG1:1096324; Let1.  
DR InterPro; IPR00775; LST1.  
DR Pfam; PF05083; LST1; 1.  
KW Immune response; Cell shape; Transmembrane;  
KW Alternative splicing.  
FT TRANSMEM 22 42  
FT VARSPLIC 53 66  
FT POTENTIAL.  
FT Missing (in isoform 2).  
FT H -> D (IN REF. 2 AND 3).  
FT Y -> N (IN REF. 2 AND 3).  
FT G -> R (IN REF. 2 AND 3).  
FT Missing (IN REF. 3).  
SQ SEQUENCE 95 AA; 10325 NW; 7B22EC72CF73C2C CRC64;  
  
Query Match 0.6%; Score 7; DB 1; Length 95;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1150 LGGLLLLL 1156  
Db 24 LGGLLLLL 30  
  
RESULT 47  
ID FEL2\_FELCA STANDARD; PRT; 109 AA.  
AC P30440;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Major allergen I polypeptide chain 2 precursor (Allergen Fel d 1-B)  
DE (Fel d 1-B) (Allergen Cat-I) (AG4) (Fdi).  
GN CH2.  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Pissipedia; Felidae; Felis.  
OX NCBI\_TaxID=9685;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-100.  
RX MEDLINE=92052157; PubMed=1946388;  
RA Morgenstern J.P., Griffith I.J., Brauer A.W., Rogers B.L.,  
Bond J.F., Chapman M.D., Kuo M.-C.;  
RT "Amino acid sequence of Fel d1, the major allergen of the domestic  
cat: protein sequence analysis and cDNA cloning,";  
EL Proc. Natl. Acad. Sci. U.S.A. 88:9690-9694(1991).  
RN [2]  
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND VARIANTS.  
RC TISSUE=Liver;  
RX MEDLINE=92241678; PubMed=1572548;  
RA Griffith I.J., Craig S., Pollock J., Yu X.-B., Morgenstern J.P.,  
Rogers B.L.;  
RT "Expression and genomic structure of the genes encoding Fd1, the  
major allergen from the domestic cat,";  
RL Gene 113:263-268(1992).  
RN [3]  
RP SEQUENCE OF 18-37, AND CHARACTERIZATION.  
RX MEDLINE=91287714; PubMed=1712068;  
RA Duffort O.A., Carreira J., Nitti G., Polo F., Lombardero M.;  
RT "Studies on the biochemical structure of the major cat allergen Felis  
domesticus I,";  
RL Mol. Immunol. 28:301-309(1991).  
RN [4]  
RP CHARACTERIZATION.  
RX MEDLINE=84265679; PubMed=6747135;  
RA Leiterman K., Ohman J.L. Jr.;

RT "Cat allergen 1: Biochemical, antigenic, and allergenic properties,";  
RL J. Allergy Clin. Immunol. 74:147-153(1984).  
CC -!- SUBUNIT: Heterotrimer composed of two non-covalently linked  
CC disulfide-linked heterodimer of chains 1 and 2.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Comment=Experimental confirmation may be lacking for some  
CC isoforms;  
CC Name=1; Synonyms=CH2L;  
CC ISCID=P30440-1; Sequence=Displayed;  
CC Name=2; Synonyms=CH2S;  
CC ISCID=P30440-2; Sequence=VSP\_004249;  
CC Name=3; Synonyms=CH2ST, Truncated;  
CC ISCID=P30440-3; Sequence=VSP\_004249;  
CC -!- TISSUE SPECIFICITY: The long form is preferentially expressed in  
CC the salivary gland, while the short form is preferentially  
CC expressed in the skin.  
CC -!- ALLERGEN: Causes an allergic reaction in human. Major allergen  
CC produced by the domestic cat.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M77341; AAC41616.1; -.  
CC EMBL; X62478; CAA44345.1; -.  
DR PIR; B53283; B53283.  
DR PIR; C56413; C56413.  
DR PIR; JCI127; JCI127.  
DR InterPro; IPR006038; Uteroglobin supf.  
KW Allergen; Glycoprotein; Signal; Polymorphism; Alternative splicing.  
FT SIGNAL 1 17  
FT CHAIN 18 109  
FT CARBOHYD 50 50  
FT VARSPLIC 82 109  
FT TISSKCD -> IAINEX (in isoform 2).  
FT /FTID=VSP\_004248.  
FT TISSKCD -> IAINEX (in isoform 2).  
FT /FTID=VSP\_004249.  
FT I -> L (IN CH2LV).  
FT I -> V (IN CH2SV).  
FT RV -> KP (IN CH2SV).  
FT M -> T (IN CH2LV).  
FT Q -> E (IN CH2SV).  
FT N -> K (IN CH2SV).  
FT C -> F (IN REF. 3).  
FT F -> T (IN REF. 3).  
SQ SEQUENCE 109 AA; 11854 MW; 857FB9CD76036CB9 CRC64;  
  
Query Match 0.6%; Score 7; DB 1; Length 109;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1157 ALLVLAL 1163  
Db 4 ALLVLAL 10  
  
RESULT 48  
HDEA\_ECOLI  
ID HDEA\_ECOLI STANDARD; PRT; 110 AA.  
AC P26604;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Protein hdeA precursor (10K-S protein).  
GN HDEA OR B3510 OR Z4922 OR ECS4390 OR SF3544 OR S4223.  
OS Escherichia coli,  
OS Escherichia coli O157:H7, and

OS Shigella flexneri.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562, 83334, 623;  
RN [1]  
RN SEQUENCE FROM N.A., AND SEQUENCE OF 22-40.  
RP SPECIES=E.coli; STRAIN=K12;  
RC MEDLINE=93204884; PubMed=8455549;  
RX Yoshida T., Ueguchi C., Yamada H., Mizuno T.;  
RA "Function of the Escherichia coli nucleoid protein, H-NS: molecular  
RT analysis of a subset of proteins whose expression is enhanced in a  
RT hns deletion mutant.";  
RL Mol. Gen. Genet. 237:113-122 (1993).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP SPECIES=E.coli; STRAIN=K12 / MG1655;  
RC MEDLINE=94316500; PubMed=8041620;  
RX Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;  
RA "Analysis of the Escherichia coli genome. V. DNA sequence of the  
RT region from 76.0 to 81.5 minutes.";  
RL Nucleic Acids Res. 22:2576-2586 (1994).  
RN [3]  
RN SEQUENCE FROM N.A.  
RP SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700952;  
RC MEDLINE=21074935; PubMed=11206551;  
RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grobeck E.J., Davis M.W., Lim A., Dimalanta E.T., Potamousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
RL Nature 409:523-533 (2001).  
RN [4]  
RN SEQUENCE FROM N.A.  
RP SPECIES=E.coli; STRAIN=O157:H7 / RMD 0509952;  
RC MEDLINE=21156231; PubMed=11258796;  
RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Yokoyama K.,  
RA Hara C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
RL DNA Res. 8:11-22 (2001).  
RN [5]  
RN GENE NAME.  
RP SPECIES=E.coli;  
RC MEDLINE=94064579; PubMed=8244952;  
RX Yoshida T., Ueguchi C., Mizuno T.;  
RA "Physical map location of a set of Escherichia coli genes (hde) whose  
RT expression is affected by the nucleoid protein H-NS.";  
RL J. Bacteriol. 175:7747-7748 (1993).  
RN [6]  
RN SEQUENCE OF 22-41.  
RP SPECIES=E.coli; STRAIN=K12 / W3110;  
RC Pasquali C., Sanchez J.-C., Ravier F., Golaz O., Hughes G.J.,  
RA Prutiger S., Paquet N., Wilkins M., Appel R.D., Bairoch A.,  
RA Hochstrasser D.F.;  
RL Submitted (SEP-1994) to Swiss-Prot.  
RN [7]  
RN SEQUENCE OF 22-33.  
RP SPECIES=E.coli; STRAIN=K12 / EMG2;  
RX MEDLINE=97443975; PubMed=9298646;  
RA Link A.J., Robison K., Church G.M.;  
RT "Comparing the predicted and observed properties of proteins encoded  
RT in the genome of Escherichia coli K-12.";  
RL Electrophoresis 18:1259-1313 (1997).  
RN [8]  
RN SEQUENCE OF 22-31.  
RP SPECIES=E.coli; STRAIN=K12;  
RC MEDLINE=99085675; PubMed=9868784;  
RX Wasinger V.C., Humphrey-Smith I.;  
RA "Small genes/gene-products in Escherichia coli K-12.";  
RT

RL RN SEQUENCE FROM N.A.  
RP SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;  
RX MEDLINE=2227406; PubMed=12384590;  
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
RA Sun L., Yue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
RA Yu J.;  
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity  
RT through comparison with genomes of Escherichia coli K12 and O157.";  
RL Nucleic Acids Res. 30:4432-4441 (2002).  
RN [10]  
RN SEQUENCE FROM N.A.  
RP SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;  
RX MEDLINE=22590274; PubMed=12704152;  
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,  
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,  
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,  
RA Schwartz D.C., Blattner F.R.;  
RT "Complete genome sequence and comparative genomics of Shigella  
RT flexneri serotype 2a strain 2457T.";  
RL Infect. Immun. 71:2775-2786 (2003).  
RN [11]  
RN MASS SPECTROMETRY.  
RP SPECIES=E.coli; STRAIN=ATCC 15597;  
RX MEDLINE=22067786; PubMed=12071744;  
RA Reid G.E., Shang H., Hogan J.M., Lee G.U., McLuckey S.A.;  
RT "Gas-phase concentration, purification, and identification of whole  
RT proteins from complex mixtures.";  
RL J. Am. Chem. Soc. 124:7353-7362 (2002).  
RN [12]  
RN X-RAY CRYSTALLOGRAPHY.  
RP SPECIES=E.coli;  
RX MEDLINE=98400493; PubMed=9731767;  
RA Yang F., Gustafson K.R., Boyd M.R., Wlodawer A.;  
RT "Crystal structure of Escherichia coli HdeA.";  
RL Nat. Struct. Biol. 5:763-764 (1998).  
RN [13]  
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RP SPECIES=E.coli;  
RX MEDLINE=20090957; PubMed=10623550;  
RA Gajiwala K.S., Burley S.K.;  
RT "HDEA, a periplasmic protein that supports acid resistance in  
RT pathogenic enteric bacteria.";  
RL J. Mol. Biol. 295:605-612 (2000).  
CC -1- SUBCELLULAR LOCATION: Periplasmic (Potential).  
CC -1- MASS SPECTROMETRY: MW=9742; METHOD=Electrospray; RANGE=22-110.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC -----  
CC EMBL: D11109; BAA01883.1; -;  
CC EMBL: U03039; AAE18486.1; -;  
CC EMBL: AB000427; AAC76835.1; -;  
CC EMBL: AB005576; AAG58851.1; -;  
CC EMBL: AP002565; BAB37813.1; -;  
CC EMBL: AE015362; AAN44999.1; -;  
CC EMBL: AE016992; AAP19187.1; -;  
CC PIR: F91177; F91177;  
CC PIR: G86023; G86023;  
CC PIR: S30268; S30268;  
CC PDB: 1BG8; 16-SEP-98.  
CC PDB: 1DJ8; 10-DEC-99.  
CC DR SWISS-2DPAGE; P26604; COLI.  
CC EcoGene; EG11398; hdeA.  
CC Periplasmic; Signal; 3D-structure; Complete proteome.  
KW

```

FT SIGNAL 1 21
FT CHAIN 22 110
FT DISULFID 39 87
FT HELIX 34 36
FT STRAND 38 38
FT HELIX 39 43
FT TURN 44 44
FT HELIX 47 49
FT HELIX 50 60
FT TURN 61 64
FT HELIX 66 68
FT HELIX 73 88
FT TURN 89 92
FT TURN 91 92
FT STRAND 94 94
FT HELIX 95 107
SQ SEQUENCE 110 AA; 11858 MW; 063262C4863FA2E9 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1150 LGGLLLL 1156
Db 9 LGGLLLL 15

RESULT 49
INS1_RAT
ID INS1_RAT STANDARD; PRT; 110 AA.
AC P01322;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Insulin 1 precursor.
GN INS1 OR INS-1
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045034; PubMed=498283;
RA Cordell B., Bell G.T., Fischer E., Denoto F.M., Ullrich A.,
Pictet R.L., Rutter W.J., Goodman H.M.;
RT "Isolation and characterization of a cloned rat insulin gene.";
RL Cell 18:533-543(1979).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; Tissue=Liver;
RX MEDLINE=80045035; PubMed=498284;
RA Lomedico P., Rosenthal N., Estratiadis A., Gilbert W., Kolodner R.,
Tizard R.;
RT "The structure and evolution of the two nonallelic rat preproinsulin
genes.";
RL Cell 18:545-558(1979).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=80240379; PubMed=6249167;
RA Lomedico P.T., Rosenthal N., Kolodner R., Estratiadis A.,
Gilbert W.;
RT "The structure of rat preproinsulin genes.";
RL Ann. N.Y. Acad. Sci. 343:425-432(1980).
RN [4]
RP SEQUENCE OF 25-54 AND 90-110.
RX MEDLINE=70067613; PubMed=4311938;
RA Steiner D.F., Clark J.U., Nolan C., Rubenstein A.H., Margoliash E.,
Aten B., Oyer P.E.;
RT "Proinsulin and the biosynthesis of insulin.";
RL Recent Prog. Horm. Res. 25:207-282(1969).
RN [5]
RP SEQUENCE OF 57-87.
RX MEDLINE=73061498; PubMed=4640931;

```

```

RA Tager H.S., Steiner D.F.;
RT "Primary structures of the proinsulin connecting peptides of the rat
and the horse.";
RL J. Biol. Chem. 247:7936-7940(1972).
RN [6]
RP SEQUENCE OF 57-87, AND REVISIONS.
RX MEDLINE=72177385; PubMed=4554104;
RA Marxhausen J., Sundby F.;
RT "Rat-proinsulin C-peptides. Amino-acid sequences.";
RL Eur. J. Biochem. 25:153-162(1972).
CC -1- FUNCTION: Insulin decreases blood glucose concentration. It
increases cell permeability to monosaccharides, amino acids and
fatty acids. It accelerates glycolysis, the pentose phosphate
cycle, and glycogen synthesis in liver.
CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
disulfide bonds.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the insulin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
DR EMBL; V01242; CA24559.1; -
DR EMBL; J00747; AAA41442.1; -
DR EMBL; M25584; AAA41439.1; -
DR PIR; A90788; IPRTI.
DR HSSP; P01308; 1A7F.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IIGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism; Signal; Multigene family.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN 1 B CHAIN.
FT PROPEP 57 87 INSULIN 1 C PEPTIDE.
FT CHAIN 90 110 INSULIN 1 A CHAIN.
FT DISULFID 31 96 INTERCHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100
SQ SEQUENCE 110 AA; 12420 MW; 51D606DA54AE3533 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1155 LLALLVL 1161
Db 10 LLALLVL 16

RESULT 50
INS_RABIT
ID INS_RABIT STANDARD; PRT; 110 AA.
AC P01311;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Insulin precursor.
GN INS.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; Tissue=Pancreas;
RX MEDLINE=94179230; PubMed=8132571;

```

```

RA Devaskar S.U., Giddings S.J., Rajakumar P.A., Carnaghi L.R.,
RA Menon R.K., Zahm D.S.;
RT "Insulin gene expression and insulin synthesis in mammalian neuronal
RT cells."; Chem. 269:8445-8454(1994).
RL J. Biol. Med. 269:8445-8454(1994).
RN [2]
RP SEQUENCE OF 25-54 AND 90-110.
RX MEDLINE=66160119; PubMed=5949593;
RA Smith L.P.;
RT "Species variation in the amino acid sequence of insulin.";
RT Am. J. Med. 40:662-666(1966).
RN [3]
RP SEQUENCE OF 56-110 FROM N.A.
RA Giddings S.J., Carnaghi L.R., Devaskar S.U.;
RA Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.
RL -!- FUNCTION: Insulin decreases blood glucose concentration. It
CC increases cell permeability to monosaccharides, amino acids and
CC fatty acids. It accelerates glycolysis, the pentose phosphate
CC cycle, and glycogen synthesis in liver.
CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
CC disulfide bonds.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the insulin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; U03610; AAA19033.1; -.
DR EMBL; M61153; AAA17540.1; -.
DR PIR; A53438; INR3.
DR HSSP; P01308; ILYM.
DR InterPro; IPR004925; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IIGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
DR Insulin family; Hormone; Glucose metabolism; Signal.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN B CHAIN.
FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 90 110 INSULIN A CHAIN.
FT DISULFID 31 96 INTERCHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100 INTERCHAIN.
FT CONFLICT 83 83 E -> Y (IN REF. 3).
SQ SEQUENCE 110 AA; 11838 MW; 82D2975B85D77FA8 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1155 LLALLVL 1161
DB 10 LLALLVL 16
|||||
|||||

RESULT 51
IGF MYXGL
ID IGF MYXGL STANDARD; PRT; 139 AA.
AC P22518;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DE Insulin-like growth factor precursor (IGF) (Fragment).
OS Myxine glutinosa (Atlantic hagfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Myxinae; Myxine.
NCBI_TaxID=7769;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91115860; PubMed=1989990;
RA Nagamatsu S., Chan S.J., Falkner S., Steiner D.F.;
RT "Evolution of the insulin gene superfamily. Sequence of a
RT preproinsulin-like growth factor cDNA from the Atlantic hagfish.";
RL J. Biol. Chem. 266:2397-2402(1991).
CC -!- FUNCTION: The insulin-like growth factors, isolated from plasma,
CC are structurally and functionally related to insulin but have a
CC much higher growth-promoting activity.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the insulin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; M57735; AAA49265.1; -.
DR PIR; A38612; A38612.
DR HSSP; P01344; IGF2.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IIGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
DR Insulin family; Mitogen; Growth factor; Signal.
FT NON TER 1 38
FT SIGNAL <1 38
FT CHAIN 39 139 INSULIN-LIKE GROWTH FACTOR.
FT DOMAIN 39 67 B.
FT DOMAIN 68 82 C.
FT DOMAIN 83 103 A.
FT DOMAIN 104 113 D.
FT DOMAIN 114 139 E.
SQ SEQUENCE 139 AA; 16087 MW; 2FC888C8D074FAC1 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 139;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLALLL 1159
DB 28 LLALLL 34
|||||
|||||

RESULT 52
LAMP PETWA
ID LAMP PETWA STANDARD; PRT; 139 AA.
AC P33575; P33576;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DB Lamprin 0.9 precursor (Cartilage matrix protein).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A. (ISOPFORMS 0.9-10 AND 0.9-12), AND SEQUENCE OF
RP 20-44.
RC TISSUE=Cartilage;
RX MEDLINE=93123269; PubMed=7678258;
RA Robson P., Wright G.M., Sitarz E., Maiti A., Rawat M., Youson J.H.,
RA Keeley F.W.;
RT "Characterization of lamprin, an unusual matrix protein from lamprey
RT cartilage. Implications for evolution, structure, and assembly of
RT elastin and other fibrillar proteins.";
RL J. Biol. Chem. 268:1440-1447(1993).
CC -!- FUNCTION: Self-aggregating protein that makes part of the soluble

```

CC form of lamprin.  
 CC -1- SUBUNIT: The polymeric lamprin chains self-aggregate to form  
 CC fibers and have secondary structures particularly rich in beta-  
 CC sheets and in beta-turns.  
 CC -1- SUBCELLULAR LOCATION: Extracellular matrix.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=0.9-12;  
 CC IsoId=P3575-1; Sequence=Displayed;  
 CC Name=0.9-10;  
 CC IsoId=P3575-2; Sequence=VSP\_004302;  
 CC Note=No experimental confirmation available;  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L05925; AAA49269.1; --  
 CC EMBL; L05924; AAA49268.1; --  
 CC PIR; C45051; C45051.  
 CC Repeat; Connective tissue; Signal; Alternative splicing.  
 CC SIGNAL 1 19  
 CC CHAIN 20 139 LAMPRIN 0.9.  
 CC DOMAIN 42 110 8 X 5 AA APPROXIMATE REPEATS.  
 CC REPEAT 42 46 1.  
 CC REPEAT 47 51 2.  
 CC REPEAT 52 56 3.  
 CC REPEAT 57 61 4.  
 CC REPEAT 62 66 5.  
 CC REPEAT 67 71 6.  
 CC REPEAT 92 96 7.  
 CC REPEAT 106 110 8.  
 CC VARSPLIC 86 104 Missing (in isoform 0.9-10).  
 CC /FTId=VSP\_004302.  
 CC SEQUENCE 139 AA; 13257 MW; B248ABB7A6CEC7C3 CRC64;  
 CC -----  
 CC Query Match 0.63; Score 7; DB 1; Length 139;  
 CC Best Local Similarity 100.0%; Pred. No. 77;  
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC -----  
 CC Qy 1157 ALLVIAL 1163  
 CC | | | | |  
 CC 7 ALLVIAL 13  
 CC -----  
 CC RESULT 53  
 CC HEAD\_CHICK STANDARD; PRT; 141 AA.  
 CC AC P02001;  
 CC DT 21-JUL-1986 (Rel. 01, Created)  
 CC DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Hemoglobin alpha-D chain.  
 CC GN HBAD.  
 CC OS Gallus gallus (Chicken).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
 CC OC Gallus.  
 CC OC NCBI TaxID=9031;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=White leghorn;  
 CC RX MEDLINE=9202023; PubMed=1656392;  
 CC RA Lewis W., Lee J.D., Dodgson J.B.;  
 CC RT "Adult chicken alpha-globin gene expression in transfected QT6 quail  
 CC cells: evidence for a negative regulatory element in the alpha D gene  
 CC region.";  
 CC RL Nucleic Acids Res. 19:5321-5329(1991).  
 CC RN [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=83161047; PubMed=6300093;  
 RA Dodgson J.B., Engel J.D.;  
 RT "The nucleotide sequence of the adult chicken alpha-globin genes.";  
 RL J. Biol. Chem. 258:4623-4629(1983).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82082384; PubMed=6273837;  
 RA Dodgson J.B., McCune K.C., Rusling D.J., Krust A., Engel J.D.;  
 RT "Adult chicken alpha-globin genes alpha A and alpha D: no anemic  
 RT shock alpha-globin exists in domestic chickens.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 78:5998-6002(1981).  
 RN [4]  
 RP PRELIMINARY SEQUENCE.  
 RX MEDLINE=76189993; PubMed=1225908;  
 RA Takei H., Ota Y., Wu K.C., Kiyohara T., Matsuda G.;  
 RT "Amino acid sequence of the alpha chain of chicken AI hemoglobin.";  
 RL J. Biochem. 77:1345-1347(1975).  
 RN [5]  
 RP SEQUENCE OF 1-63 AND 94-122.  
 RC TISSUE=Embryo;  
 RX MEDLINE=82098109; PubMed=7054172;  
 RA Chapman B.S., Hood L.E., Tobin A.J.;  
 RT "Minor early embryonic chick hemoglobin M. Amino acid sequences of  
 RT the epsilon and alpha D chains.";  
 RL J. Biol. Chem. 257:651-658(1982).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=99156923; PubMed=10037733;  
 RA Knapp J.E., Oliveira M.A., Xie Q., Ernst S.R., Riggs A.F.,  
 RA Hackert M.B.;  
 RT "The structural and functional analysis of the hemoglobin D component  
 RT from chicken.";  
 RL J. Biol. Chem. 274:6411-6420(1999).  
 CC -1- FUNCTION: Involved in oxygen transport from the lung to the  
 CC various peripheral tissues.  
 CC -1- SUBUNIT: Heterotetramer of two alpha-D chains and two beta chains.  
 CC The component D forms dimers of tetramers upon deoxygenation.  
 CC -1- TISSUE SPECIFICITY: Red blood cells.  
 CC -1- DEVELOPMENTAL STAGE: In birds, the alpha-D chain occurs in a minor  
 CC hemoglobin component, called hemoglobin d, which is expressed in  
 CC late embryonic and adult life.  
 CC -1- MISCELLANEOUS: REP.5 CHAIN WAS ISOLATED FROM HEM, THE LEAST  
 CC ABUNDANT OF THE FOUR EARLY CHICK HEMOGLOBINS.  
 CC -1- SIMILARITY: Belongs to the globin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X59889; CAA42605.1; --  
 CC EMBL; V00411; CAA23702.1; --  
 CC EMBL; M15378; AAA48584.1; ALT\_SEQ.  
 CC EMBL; J00853; AAA48800.1; --  
 CC PIR; B92421; EACHI.  
 CC PDB; 1HBR; 26-MAR-99.  
 CC InterPro; IPR002338; Alpha.haem.  
 CC InterPro; IPR000971; Globin.  
 CC Pfam; PF00042; Globin; 1.  
 CC PRINTS; PR00612; ALPHAHAE.  
 CC PROSITE; PS01033; GLOBIN; 1.  
 CC Heme; Oxygen transport; Transport; Erythrocyte;  
 CC 3D-structure.  
 CC METAL 58 58 IRON (HEME DISTAL LIGAND).  
 CC METAL 87 87 IRON (HEME PROXIMAL LIGAND).  
 CC CONFLICT 16 16 K -> R (IN REF. 1).  
 CC CONFLICT 107 107 V -> C (IN REF. 3 AND 5).  
 CC HELIX 4 17  
 CC TURN 18 19

```

FT HELIX      21      35
FT HELIX      37      42
FT TURN       44      45
FT TURN       50      51
FT HELIX      53      71
FT TURN       72      74
FT HELIX      76      79
FT TURN       81      88
FT HELIX      89      91
FT TURN       95      95
FT HELIX      96      113
FT HELIX     114      116
FT HELIX     119      136
FT TURN     137      137
SQ SEQUENCE 141 AA; 15695 MW; 1FE426969B7B5384 CRC64;

Query Match      0.6%; Score 7; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY      625 GAGNAV 631
Db      |||||
        64 GAGNAV 70

RESULT 54
HEAD MELGA      STANDARD;      PRT;      141 AA.
AC F81024;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hemoglobin alpha-D chain.
GN HBAD.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE.

MEDLINE=96066279; PubMed=7576241;
RA Eguchi Y., Ikehara T., Kayo S., Eguchi T., Takei H.;
RT "Amino acid sequence of alpha- and beta-polypeptide chains of turkey
  (Meleagris gallopavo) hemoglobin.";
RL Biol. Chem. Hoppe-Seyler 376:437-440(1995).
CC -!- FUNCTION: Involved in oxygen transport from the lung to the
  various peripheral tissues.
CC -!- SUBUNIT: Heterotetramer of two alpha-D chains and two beta chains.
CC -!- TISSUE SPECIFICITY: Red blood cells.
CC -!- DEVELOPMENTAL STAGE: In birds, the alpha-D chain occurs in a minor
  late embryonic component, called hemoglobin d, which is expressed in
  late embryonic and adult life.
CC -!- SIMILARITY: Belongs to the globin family.
PIR; S56103; S56103.
DR HSSP; P02001; IHRB.
DR InterPro; IPR002338; Alpha_haem.
DR InterPro; IPR000971; Globin.
DR Pfam; PF00042; globin; 1.
DR PRINTS; PR00612; ALPHAHAE.
DR PROSITE; PS01033; GLOBIN; 1.
KW Heme; Oxygen transport; Transport; Erythrocyte.
FT METAL      58      58
FT METAL      87      87
FT METAL      87      87
SQ SEQUENCE 141 AA; 15665 MW; 59942696887954E7 CRC64;

Query Match      0.6%; Score 7; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      625 GAGNAV 631
Db      |||||
        64 GAGNAV 70

```

```

RESULT 55
MK_CHICK
ID MK_CHICK      STANDARD;      PRT;      142 AA.
AC P24052;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Midkine precursor (Retinoic acid-induced heparin-binding protein)
  (RI-HB).
DE (RI-HB).
GN RIHB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 80-85; 112-118 AND 121-133.
RC TISSUE=Embryo;
RX MEDLINE=91207359; PubMed=2018506;
RA Urios P., Duprez D., le Caer J.-P., Courtois Y., Vigny M., Laurent M.;
RT "Molecular cloning of RI-HB, a heparin binding protein regulated by
  retinoic acid.";
RT retinoic acid.";
RL Biochem. Biophys. Res. Commun. 175:617-624(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95010085; PubMed=7925417;
RA Duprez D., Treaggar J., Pecqueur C., Vigny M.R.;
RT "Organisation and promoter activity of the
  retinoic-acid-induced-heparin-binding (RIHB) gene.";
RL Eur. J. Biochem. 224:931-941(1994).
RN [3]
RP SEQUENCE OF 22-77.
RX MEDLINE=91128406; PubMed=1993066;
RA Raulais D., Lagente-Chevallier O., Guettet C., Duprez D., Courtois Y.,
  Vigny M.;
RT "A new heparin binding protein regulated by retinoic acid from chick
  embryo.";
RL Biochem. Biophys. Res. Commun. 174:708-715(1991).
RN [4]
RP SEQUENCE OF 22-77.
RC STRAIN=White leghorn; TISSUE=Embryo;
RX MEDLINE=90108010; PubMed=2558016;
RA Vigny M., Raulais D., Puzenat N., Duprez D., Hartman M.P.,
  Jeanny J.C., Courtois Y.;
RT "Identification of a new heparin-binding protein localized within
  chick basement membranes.";
RL Eur. J. Biochem. 186:733-740(1989).
CC -!- FUNCTION: Has mitogenic activity, and neurite extension activity
  for PC12 cells.
CC -!- SUBCELLULAR LOCATION: BASEMENT MEMBRANES IN EARLY EMBRYONIC
  TISSUES, AND CELL SURFACE OF NEUROECTODERMAL CELLS.
CC -!- DEVELOPMENTAL STAGE: Essentially expressed during embryogenesis.
CC -!- INDUCTION: By retinoic acid.
CC -!- SIMILARITY: Belongs to the pleiotrophin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
DR EMBL; M61754; -; NOT ANNOTATED_CDS.
DR EMBL; X76482; CAA54020.1; -.
DR PIR; J70573; J70573.
DR HSSP; P21741; LMKC.
DR InterPro; IPR000762; PTN_MK.
DR Pfam; PF01091; PTN_MK_C; 1.
DR Pfam; PF05196; PTN_MK_N; 1.
DR PRINTS; PR00269; PTNMIDKINE.
DR ProDom; PD005592; PTN_MK; 1.

```

DR SMART: SM00193; PTN: 1;  
DR PROSITE; PS00619; PTN MK 1; 1;  
DR PROSITE; PS00620; PTN MK 2; 1;  
KW Growth factor; Mitogen; Differentiation; Heparin-binding; Signal.  
FT SIGNAL 1 21  
FT CHAIN 22 142 MIDKINE.  
FT DISULFID 36 60 BY SIMILARITY.  
FT DISULFID 44 69 BY SIMILARITY.  
FT DISULFID 51 73 BY SIMILARITY.  
FT DISULFID 83 115 BY SIMILARITY.  
FT DISULFID 93 125 BY SIMILARITY.  
FT CONFLICT 89 89 S -> R (IN REF. 2).  
FT CONFLICT 93 93 C -> G (IN REF. 2).  
SQ SEQUENCE 142 AA; 15579 MW; 9D05CAF89558451B CRC64;  
  
Query Match 0.68; Score 7; DB 1; Length 142;  
Best Local Similarity 100.0%; Pred. No. 78; 0; Caps 0;  
Matches 7; Conservative 0; Mismatches 0; Indels 0;  
  
Qy 1153 LLLALL 1159  
Db 7 LLLALL 13  
  
RESULT 56  
BIK HUMAN  
ID \_BIK\_HUMAN STANDARD; PRT; 160 AA.  
AC Q13523; Q16582;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Bcl-2 interacting killer (Apoptosis inducer NBK) (BK4) (BIPI).  
GN BIK OR NBK.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutarchia; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=B-cell;  
RX MEDLINE=96068922; PubMed=7478623;  
RA Boyd J.M., Gallo G.J., Elangovan B., Houghton A.B., Malstrom S.,  
RA Avery B.J., Ebb R.G., Subramanian T., Chittenden T., Lutz R.J.,  
RA Chinnadurai G.;  
RT "Bik, a novel death-inducing protein shares a distinct sequence motif  
RT with Bcl-2 family proteins and interacts with viral and cellular  
RT survival-promoting proteins."  
RL Oncogene 11:1921-1928(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9641339; PubMed=8816500;  
RA Han J., Sabbatini P., White E.;  
RT "Induction of apoptosis by human NBK/Bik, a BH3-containing protein  
RT that interacts with E1B 19K."  
RL Mol. Cell. Biol. 16:5857-5864(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymphoid;  
RA Pun K.-T., Farrow S.N., Raven T., Wride C.J., White J.H.M., Brown R.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99431905; PubMed=10500065;  
RA Castellani A., Ino Y., Louis D.N., Ramesh V., Gusella J.F., Rustgi A.K.;  
RT "Mapping of a target region of allelic loss to a 0.5-cM interval on  
RT chromosome 22q13 in human colorectal cancer."  
RL Gastroenterology 117:831-837(1999).  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Rieder M.J., Livingston R.J., Daniels M.R., Chung M.-W.,  
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,  
RA Schackwitz W.S., Sherwood J.K., Wittrak L.A., Nickerson D.A.;  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.

[6]  
SEQUENCE FROM N.A.  
MEDLINE=20057165; PubMed=10591208;  
RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,  
RA Clump M., Smink L.J., Ainscough R., Almeida J.P., Babage A.K.,  
RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,  
RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,  
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,  
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,  
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,  
RA Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,  
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A., Hall C.,  
RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,  
RA Hunt R.E., Hall-Tamlyn G., Heathcote S.W., Ho S., Holmes S.,  
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,  
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,  
RA Marjyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,  
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,  
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,  
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,  
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,  
RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,  
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,  
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,  
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudon J.,  
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,  
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,  
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,  
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,  
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,  
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,  
RA Zhan M., Zhang G., Chisoe S., Murray J.J., Miller N., Minx P.,  
RA Fulton R., Johnson D., Hemis G., Bentley D., Bradshaw H., Bourne S.,  
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,  
RA Hinds K., Kemp K., Latreille P., Layman D., Ozerky P., Rohlfing T.,  
RA Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,  
RA Korf I., Bedell J.A., Hillier L.M., Mardis E., Waterston R.,  
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,  
RA Sudar M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,  
RA Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,  
RA Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,  
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,  
RA Khan A.S., Lane L., Tilahun Y., Wright H.;  
RT "The DNA sequence of human chromosome 22."  
RL Nature 402:489-495(1999).  
RN [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymph;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Rutherford Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [8]  
RP MUTAGENESIS, AND FUNCTION OF BH3 DOMAIN.  
RX MUTAGENESIS, AND FUNCTION OF BH3 DOMAIN.  
MEDLINE=96091131; PubMed=8521816;



```

RA Chittenden T., Flemington C., Houghton A.B., Ebb R.G., Gallo G.J.,
RA Elangovan B., Chinnadurai G., Lutz R.J.;
RT "A conserved domain in Bak, distinct from BH1 and BH2, mediates cell
RT death and protein binding functions.";
RL EMBO J. 14:5589-5596(1995)
CC -!- FUNCTION: Accelerates programmed cell death. Binding to the
CC apoptosis repressors Bcl-X(L), BHRF1, Bcl-2 or its adenovirus
CC homolog E1B 19k protein suppresses this death-promoting activity.
CC Does not interact with BAX.
CC -!- SUBCELLULAR LOCATION: Around the nuclear envelope, and in
CC cytoplasmic membranes.
CC -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD and
CC BAX for their pro-apoptotic activity and for their interaction
CC with anti-apoptotic members of the Bcl-2 family.
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U34584; AAC50413.1; -
DR EMBL; U49730; AAC79124.1; -
DR EMBL; X89986; CA862013.1; -
DR EMBL; AF174424; AAF01156.1; -
DR EMBL; AF174421; AAF01156.1; JOINED.
DR EMBL; AF174422; AAF01156.1; JOINED.
DR EMBL; AF174423; AAF01156.1; JOINED.
DR EMBL; AY245248; AAO61089.1; -
DR EMBL; AL022237; CAA18260.2; -
DR EMBL; BC001599; AAO01599.1; -
DR EMBL; S89214; S58214.
DR EMBL; BC001599; AAO01599.1; -
DR EMBL; HGNC:1051; BIK.
DR MIM; 603392; -
DR GO; GO:0008632; P:apoptotic program; TAS.
DR GO; GO:0006917; P:induction of apoptosis; TAS.
DR InterPro; IPR000712; Bcl2_BH.
DR PROSITE; PS01259; BH3; 1.
KW Apoptosis; Transmembrane.
FT DOMAIN 57 71 BH3.
FT TRANSMEM 136 156 POTENTIAL.
FT DOMAIN 137 158 LEUCINE-ZIPPER (POTENTIAL).
FT CONFLICT 149 150 PL -> LP (IN REF. 1).
SQ SEQUENCE 160 AA; 18016 MW; 89034F443F5A136 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1153 LLLLLL 1159
DB 141 LLLLLL 147
|||||

RESULT 57
NUOE AQUAE STANDARD; PRT; 160 AA.
AC O66842.
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NADH-quinone oxidoreductase chain E (EC 1.6.99.5) (NADH dehydrogenase
DE I, chain E) (NDH-1, chain E).
GN NUOE OR AQ 574.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;

```

```

RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.B., Overbeek R., Snead M.A., Keller M., Aujaay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -!- FUNCTION: NADH-1 shuttles electrons from NADH, via FMN and iron-
CC sulfur (Fe-S) clusters, to quinones in the respiratory chain.
CC Couples the redox reaction to proton translocation (for every two
CC electrons transferred, four hydrogen ions are translocated across
CC the cytoplasmic membrane), and thus conserves the redox energy in
CC a proton gradient (by similarity).
CC -!- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -!- COFACTOR: Binds 1 2Fe-2S cluster (Potential).
CC -!- SIMILARITY: Belongs to the complex I 24 kDa subunit family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000696; AAC06799.1; -
DR EMBL; PIR; P70351; P70351.
DR InterPro; IPR02023; Cmplx1_24kDa.
DR Pfam; PF01257; complex1_24kDa; 1.
DR ProDom; PD003859; Cmplx1_24kDa; 1.
DR PROSITE; PS01099; COMPLEX1_24K; 1.
KW Oxidoreductase; NAD; Quinone; Metal-binding; Iron-sulfur; Iron;
FT METAL 86 86 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 91 91 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 127 127 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 131 131 IRON-SULFUR (2FE-2S) (POTENTIAL).
SQ SEQUENCE 160 AA; 18550 MW; 425D81995A491B3B CRC64;

Query Match 0.6%; Score 7; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 413 EFPEELK 419
DB 7 EFPEELK 13
|||||

RESULT 58
TCH2 ARATH STANDARD; PRT; 161 AA.
AC P25070; O22592;
DT 01-MAY-1992 (Rel. 22, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Calmodulin-related protein 2, touch-induced.
GN TCH2 OR AT5G3770 OR K22F20.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Johnson K.A., Braam J.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=99087489; PubMed=987454;
RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
RA Tabata S.;

```

RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.  
RT Sequence features of the regions of 1,013,767 bp covered by sixteen  
RT physically assigned P1 and TAC clones."  
RL DNA Res. 5:297-308(1998).  
RN [3]  
RP SEQUENCE OF 26-70 FROM N.A.  
RC STRAIN=CV, Columbia;  
RX MEDLINE=90150263; PubMed=2302732;  
RA Braam J., Davis R.W.;  
RT "Rain-, wind-, and touch-induced expression of calmodulin and  
RT calmodulin-related genes in Arabidopsis."  
RL Cell 60:357-364(1990).  
RN [4]  
RP 3D-STRUCTURE MODELING OF 7-158.  
RX MEDLINE=97189489; PubMed=9037719;  
RA Zhan A.R., Johnson K.A., Braam J., James M.N.G.;  
RT "Comparative modeling of the three-dimensional structure of the  
RT calmodulin-related TCH2 protein from Arabidopsis."  
RL Proteins 27:144-153(1997).  
CC -!- INDUCTION: By rain-, wind-, and touch (thigmomorphogenesis).  
CC -!- SIMILARITY: Contains 4 EF-hand calcium-binding domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; AB026473; AB82713.1; -  
DR EMBL; AB016873; BAB10353.1; -  
DR PIR; C34669; C34669.  
DR PDB; 1AVJ; 28-JAN-98.  
DR InterPro; IPR003299; Calflagin.  
DR InterPro; IPR002048; EF-hand.  
DR Pfam; PF00036; ehand; 4  
DR PRINTS; PR01362; CALFLAGIN.  
DR PRODOM; PD000012; EF-hand; 2.  
DR SMART; SM00054; EFh; 4.  
DR PROSITE; PS00018; EF HAND; 4.  
KW Calcium-binding; Repeat; 3D-structure.  
FT CA\_BIND 26 37 EF-HAND 1 (POTENTIAL).  
FT CA\_BIND 62 73 EF-HAND 2 (POTENTIAL).  
FT CA\_BIND 103 114 EF-HAND 3 (POTENTIAL).  
FT CA\_BIND 139 150 EF-HAND 4 (POTENTIAL).  
FT CONFLICT 54 54 T -> Y (IN REF. 3).  
FT HELIX 12 25  
FT TURN 27 28  
FT STRAND 33 33  
FT HELIX 35 44  
FT TURN 45 45  
FT HELIX 51 61  
FT TURN 63 64  
FT STRAND 69 69  
FT HELIX 71 102  
FT TURN 104 105  
FT STRAND 110 110  
FT HELIX 112 121  
FT TURN 122 123  
FT HELIX 128 138  
FT STRAND 146 146  
FT HELIX 148 156  
SQ SEQUENCE 161 AA; 17546 MW; A7324A5C0B3CDB9 CRC64;  
  
Query Match 0.6%; Score 7; DB 1; Length 161;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 487 VDIDGDG 493  
DB 138 VDIDGDG 144

RESULT 59  
YF87 METJA STANDARD; PRT; 171 AA.  
AC Q58982;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein MJ1587.  
GN MJ1587.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcales; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=868087;  
RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kerlavage A.R., Dougherty B.A., Tomb J.-P., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghegan N.S.M., Weidman J.D., Sadow P.W., Borodovsky M.,  
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Borodovsky M.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
RT jannaschii."  
RL Science 273:1058-1073(1996).  
CC -!- SIMILARITY: TO M.JANNASCHII MJ0417.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; U67599; AAB99616.1; -  
DR PIR; B64498; B64498.  
DR TIGR; MJ1587; -  
DR InterPro; IPR003141; PHP N.  
DR Pfam; PF02231; PHP N; 1.  
DR SMART; SM00481; POLIILAC; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 171 AA; 19198 MW; 7D4F7B936C5C7472 CRC64;  
  
Query Match 0.6%; Score 7; DB 1; Length 171;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 344 ALGDRIF 350  
DB 123 ALGDRIF 129  
  
RESULT 60  
YD2F SCHPO STANDARD; PRT; 176 AA.  
AC Q10263;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Very hypothetical protein C56F8.15 in chromosome I.  
GN SPAC5F8.15.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomyces.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.

```

RC STRAIN=972;
RX MEDLINE=21048401; PubMed=11895360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feitwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle Z.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Summers M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymoprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borszym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle S.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J.L., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z69728; CAA93586.1; -.
DR PIR; T38925; T38925.
DR GeneDB; SPAC56F8.15; -.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 27 47 POTENTIAL.
FT TRANSMEM 79 99 POTENTIAL.
FT TRANSMEM 107 127 POTENTIAL.
SQ SEQUENCE 176 AA; 20595 MW; B0FB48B199DA622A CRC64;

Query Match 0.6%; Score 7; DB 1; Length 176;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 820 YTLSPDT 826
DB 138 YTLSPDT 144

RESULT 61
ID HIS7 CHLTF STANDARD; PRT; 194 AA.
AC Q8KEF4.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19) (IGPD).
GN HISB OR CT0735.
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901;

Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.I., Yang P.,
Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
Nierman W.C., Feldblyum T.V., Hansen C.D., Craven M.B., Radune D.,
Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
Natarajan V., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium."
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
CC -!- CATALYTIC ACTIVITY: D-erythro-1-(imidazol-4-yl)glycerol 3-
CC phosphate = 3-(imidazol-4-yl)-2-oxopropyl phosphate + H(2)O.
CC -!- PATHWAY: Histidine biosynthesis; sixth step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -!- SIMILARITY: Belongs to the imidazoleglycerol-phosphate dehydratase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB012844; ANM71972.1; -.
DR TIGR; CT0735; -.
DR HAWAP; MF 00076; -.
DR InterPro; IPR000807; IGPD.
DR Pfam; PF00475; IGPD; 1.
DR ProDom; PD002282; IGPD; 1.
DR PROSITE; PS00954; IGP_DEHYDRATASE_1; 1.
DR PROSITE; PS00955; IGP_DEHYDRATASE_2; 1.
KW Histidine biosynthesis; lyase; Complete proteome.
SQ SEQUENCE 194 AA; 21189 MW; D87295AAC9B830F2 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 194;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 IVDALGD 347
DB 80 IVDALGD 86

RESULT 62
ID Z239_MOUSE STANDARD; PRT; 201 AA.
AC P24399;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 239 (Zfp-239) (Zinc finger protein MOK-2).
GN ZNF239 OR ZFP239 OR MOK2 OR MOK-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9007859; PubMed=2104662;
RA Enrout-Lange M., Kress M., Hamer D.;
RT "A gene that encodes a protein consisting solely of zinc finger
RT domains is preferentially expressed in transformed mouse cells."
RL Mol. Cell. Biol. 10:418-421(1990).
CC -!- FUNCTION: May function as a transcription factor.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN TRANSFORMED
CC MOUSE CELLS.
CC -!- SIMILARITY: BELONGS TO THE KRUPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: Contains 7 C2H2-type zinc fingers.
CC -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; M32057; AAA39729.1; -;  
 DR PIR; I57505;  
 DR HSSP; P08047; I5P2.  
 DR TRANSFAC; T00510; -;  
 DR MGD; MGI:1306812; Zfp239.  
 DR InterPro; IPR007087; Znf\_C2H2.  
 DR InterPro; IPR007086; Znf\_C2H2\_sub.  
 DR Pfam; PF00096; zf-C2H2; 7.  
 DR PRINTS; PR00048; ZINC\_FINGER.  
 DR PRODOM; PD000003; Znf\_C2H2; 6.  
 DR SMART; SM00355; Znf\_C2H2; 7.  
 DR PROSITE; PS00029; ZINC\_FINGER\_C2H2\_1; 7.  
 DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 7.  
 KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;  
 KW Nuclear protein; Repeat.  
 FT ZN\_FING 6 28 C2H2-TYPE 1.  
 FT ZN\_FING 34 56 C2H2-TYPE 2.  
 FT ZN\_FING 62 84 C2H2-TYPE 3.  
 FT ZN\_FING 90 112 C2H2-TYPE 4.  
 FT ZN\_FING 118 140 C2H2-TYPE 5.  
 FT ZN\_FING 146 168 C2H2-TYPE 6.  
 FT ZN\_FING 174 196 C2H2-TYPE 7.  
 SQ SEQUENCE 201 AA; 22832 MW; E7A18COAADAB1384 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 959 FTRESSL 965  
 DB 15 FTRESSL 21

## RESULT 63

ID GPBB\_MOUSE STANDARD; PRT; 206 AA.  
 AC P56400;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Platelet glycoprotein Ib beta chain precursor (GP-Ib beta) (GP1BB)  
 DE (GP1b-beta).  
 GN GP1BB.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97403789; PubMed=9259114;  
 RA Kitaguchi T., Murata M., Anbo H., Moriki T., Ikeda Y.;  
 RT "Characterization of the gene encoding mouse platelet glycoprotein Ib  
 RT beta.";  
 RL Thromb. Res. 87:235-244 (1997).  
 CC -!- FUNCTION: GP-Ib, a surface membrane protein of platelets,  
 CC participates in the formation of platelet plugs by binding to von  
 CC Willebrand factor, which is already bound to the subendothelium  
 CC (by similarity).  
 CC -!- SUBUNIT: GP-Ib alpha and beta are disulfide linked. GP-IX is  
 CC complexed with the GP-Ib heterodimer via a non covalent linkage  
 CC (by similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- MISCELLANEOUS: Platelet activation apparently involves disruption  
 CC of the macromolecular complex of GP-Ib with the platelet  
 CC glycoprotein IX (GP-IX) and dissociation of GP-Ib from the actin-

CC binding protein.  
 CC -!- SIMILARITY: Contains 1 leucine-rich (LRR) repeat.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; AB001419; BAA22424.1; -;  
 DR MGD; MGI:107852; Gp1bb.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000483; LRR\_Cterm.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR Pfam; PF00560; LRR; 1.  
 DR Pfam; PF01463; LRRCT; 1.  
 DR Pfam; PF01462; LRRNT; 1.  
 DR SMART; SM00082; LRRCT; 1.  
 DR SMART; SM00013; LRRNT; 1.  
 KW Platelet; Transmembrane; Glycoprotein; Hemostasis; Blood coagulation;  
 KW Signal; Phosphorylation; Cell adhesion; Leucine-rich repeat.  
 FT SIGNAL 1 26 BY SIMILARITY.  
 FT CHAIN 27 206 PLATELET GLYCOPROTEIN IB BETA CHAIN.  
 FT DOMAIN 27 147 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 148 172 POTENTIAL.  
 FT DOMAIN 173 206 CYTOPLASMIC (POTENTIAL).  
 FT REPEAT 60 83 LRR.  
 FT CARBOHYD 66 66 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT MOD RES 191 191 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
 SQ SEQUENCE 206 AA; 21762 MW; AC4BCB4DFA226FID CRC64;

Query Match 0.6%; Score 7; DB 1; Length 206;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALL 1159  
 DB 13 LLLALL 19

## RESULT 64

PSMB\_SULSO STANDARD; PRT; 208 AA.  
 ID PSMB\_SULSO  
 AC Q9UXF3;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Proteasome beta subunit precursor (EC 3.4.25.1) (Multicatalytic  
 DE endopeptidase complex beta subunit).  
 GN PSMB OR SSO0766 OR C40\_002.  
 OS Sulfolobus solfataricus.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OC Sulfolobus.  
 OX NCBI\_TaxID=2287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ATCC 35092 / DSM 1617 / P2;  
 RX MEDLINE=20165948; PubMed=10701121;  
 RA Charlebois R.L., Singh R.K., Chan-Weher C.C.-Y., Allard G., Chow C.,  
 RA Confalonieri T., Curtis B., Duguet M., Brauso G., Faguy D.,  
 RA Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,  
 RA Kuehwa N., Lafleur E., Medina N., Peng X., Penny S.L., She Q.,  
 RA St Jean A., van der Oost J., Young P., Zivanovic Y., Doolittle W.P.,  
 RA Ragan M.A., Sengen C.W.;  
 RT "Gene content and organization of a 281-kbp contig from the genome of  
 RT the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";  
 RN Genome 43:116-136(2000).  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ATCC 35092 / DSM 1617 / P2;  
 RX MEDLINE=21332296; PubMed=11427726;

```

RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweez M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Reader P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
RL The complete genome of the crenarchaeon Sulfolobus solfataricus P2.*;
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -!- FUNCTION: The proteasome is a multicatalytic proteinase complex
CC which is characterized by its ability to cleave peptides with Arg,
CC Phe, Tyr, Leu, and Glu adjacent to the leaving group at neutral or
CC slightly basic pH (By similarity).
CC -!- CATALYTIC ACTIVITY: Cleavage at peptide bonds with very broad
CC specificity.
CC -!- SUBUNIT: Composed of two subunits, alpha and beta. The complex is
CC formed of four rings. The two outer rings are each composed of
CC seven alpha subunits. The two inner rings are each composed of
CC seven beta subunits (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to peptidase family T1B.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y18930; CAB57537.1; ALT INIT.
CC EMBL; A5006700; AAK41063.1; ALT_INIT.
CC HSSP; P28061; LPMA.
CC MEROPS; T01.002; -.
CC InterPro; IPR000243; Pept_T1A_subb.
CC InterPro; IPR001353; Peptidase_T1.
CC Pfam; PF00227; Proteasome; 1.
CC PRINTS; PR00141; PROTEASOME.
CC PROSITE; PS00854; PROTEASOME B; 1.
CC Proteasome; Hydrolase; Protease; Complete proteome;
KW Threonine protease.
FT PROPEP 1 14 APPARENTLY REMOVED IN MATURE FORM (BY
FT SIMILARITY).
FT CHAIN 15 208 PROTEASOME BETA SUBUNIT.
FT ACT SITE 15 15 PROBABLE.
SQ SEQUENCE 208 AA; 22927 MW; C940A8273AB25CB9 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 686 TTTVGIR 692
DB 15 TTTVGIR 21

RESULT 65
KPTA_PYRAE STANDARD; PRT; 213 AA.
AC Q8ZSP2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable RNA 2'-phosphotransferase (EC 2.7.7.-).
GN KPTA OR PAE3647.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;

```

```

RA Fitz-Gibbon S.F., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum."
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
CC -!- FUNCTION: Removes the 2'-phosphate from RNA via an intermediate in
CC which the phosphate is ADP-ribosylated by NAD followed by a
CC presumed transesterification to release the RNA and generate ADP-
CC ribose 1',2'-cyclic phosphate (APPR>P). May function as an ADP-
CC ribosylase (By similarity).
CC -!- SIMILARITY: Belongs to the kptA / TPT1 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE009943; AAL65071.1; -.
CC HAMAP; MF_00299; -.
CC InterPro; IPR002745; Ptrans_KptA/Tpt1.
CC Pfam; PF01885; PTS_2-RNA; 1.
CC ProDom; PD009619; P04transfrs; 1.
CC Transfrase; NAD; Complete proteome.
SQ SEQUENCE 213 AA; 23452 MW; 12DF2AD12CE69636 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 DVYKCPV 78
DB 3 DVYKCPV 9

RESULT 56
CH13_HUMAN STANDARD; PRT; 214 AA.
AC Q9K5S9; Q9N3M3; Q9NSR0.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C8orf13.
GN C8ORF13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Appel S., Bergheim A., Reichwald K., Reis A., Rosenthal A., Ramsay M.,
RA Hennies H.;
RT "Transcript map of the KWE critical region on chromosome 8p22-p23."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Amalgam;
RA Wambutt R., Heubner D., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 83-214 FROM N.A.
RC TISSUE=Amalgam;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SEC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ301564; CAC82740.1; -  
DR EMBL; AL834122; CAD38843.1; -  
DR EMBL; AL157475; CAB75670.1; -  
DR PIR; T46905; T46905.  
DR Genew; HGNC:15549; C8orf13.  
KW Hypothetical protein; Coiled coil.  
FT DOMAIN 118 156 COILED COIL (POTENTIAL).  
FT CONFLICT 56 56 H -> Q (IN REF. 2).  
FT CONFLICT 107 107 T -> S (IN REF. 3).  
FT CONFLICT 130 130 M -> L (IN REF. 2).  
SQ SEQUENCE 214 AA; 24182 MW; 14216458987284 CRC64;  
  
Query Match 0.6%; Score 7; DB 1; Length 214;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1090 LRSUKAL 1096  
Db 27 LRSUKAL 33  
  
RESULT 67  
CYB\_CERCE STANDARD; PRT; 214 AA.  
AC P87419;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cytochrome b (Fragment).  
GN MTCYB OR COB OR CYTB.  
OS Cerastes cerastes (Horned desert viper).  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;  
OC Viperidae; Viperinae; Cerastes.  
OX NCBI\_TaxID=8697;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98334559; PubMed=9667984;  
ZA Vidal N., Lecointre G., Vie J.-C., Gasc J.-P.;  
RT "Weighting and congruence: a case study based on three mitochondrial genes in pitvipers.";  
RL Mol. Phylogenet. Evol. 9:366-374(1998).  
RN [2]  
RP SEQUENCE OF 1-132 FROM N.A.  
RA Vidal N., Lecointre G., Vie J.-C., Gasc J.-P.;  
RT "Molecular systematics of pitvipers: paraphyly of the Bothrops complex.";  
RL C. R. Acad. Sci., III, Sci. Vie 320:95-101(1997).  
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis (By similarity).  
CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or b562) is low-potential and absorbs at about 562, and heme 2 (or BH or b566) is high-potential and absorbs at about 566 (By similarity).  
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b, cytochrome c1 and the Rieske protein (By similarity).  
CC -!- SIMILARITY: Belongs to the cytochrome b family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF039265; AAC33542.1; -

DR InterPro; IPR005798; Cytb\_b6\_C.  
DR InterPro; IPR005797; Cytb\_b6\_N.  
DR Pfam; PF00033; Cytochrome b N; 1.  
DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; FALSE NEG.  
DR PROSITE; PS00193; CYTOCHROME\_B\_CO; PARTIAL.  
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;  
KW Heme.  
FT NON\_TER 1 1  
FT METAL 81 81 IRON 1 (HEME B562 AXIAL LIGAND).  
FT METAL 95 95 IRON 2 (HEME B566 AXIAL LIGAND).  
FT NON\_TER 214 214  
SQ SEQUENCE 214 AA; 24230 MW; 755818DB9204P820 CRC64;  
  
Query Match 0.6%; Score 7; DB 1; Length 214;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 37 RTAFPCY 43  
Db 123 RTAFPCY 129  
  
RESULT 68  
FGF3\_CHICK STANDARD; PRT; 220 AA.  
AC P48801;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fibroblast growth factor-3 precursor (FGF-3) (HBGF-3).  
GN FGF3 OR FGF-3.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Rhode Island red; TISSUE=Embryo;  
RX MEDLINE=95309122; PubMed=7789270;  
RA Mahmood R., Kiefer P., Guthrie S., Dickson C., Mason I.;  
RT "Multiple roles for FGF-3 during cranial neural development in the chicken.";  
RL Development 121:1399-1410(1995).  
CC -!- FUNCTION: POTENT MITOGEN AND TRANSFORMING AGENT (BY SIMILARITY).  
CC -!- SIMILARITY: Belongs to the heparin-binding growth factors family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z47555; CAA87635.1; -  
DR PIR; I50588; I50588.  
DR HSP; P31371; IG82.  
DR InterPro; IPR008996; Cytok\_IL1\_like.  
DR InterPro; IPR002348; IL1\_HBGF.  
DR Pfam; PF00167; FGF; 1.  
DR PRINTS; PR00262; IL1HBGF.  
DR PRODOM; PD000831; IL1\_HBGF; 1.  
DR SMART; SM00442; FGF; 1.  
DR PROSITE; PS00247; HBGF\_FGF; 1.  
KW Growth factor; Mitogen; Signal; Glycoprotein.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 220 FIBROBLAST GROWTH FACTOR-3.  
FT CARBOHYD 66 66 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 220 AA; 25050 MW; B15D41D1E551C5D5 CRC64;  
  
Query Match 0.6%; Score 7; DB 1; Length 220;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;





```

CC EMBL; L45768; AAA83027.1; -.
DR HSP; P13727; IHB0.
DR MG; WGI.103294; PrG2.
DR InterPro; IPR002352; Emaior_basic.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR PRINTS; PR00770; EMAJORBASIC.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN 1; FALSE_NEG.
DR Eosinophil; Signal; Immune response; Antibiotic; Lectin.
FT SIGNAL 1 16
FT PROPEP 17 106
FT CHAIN 107 223
FT DOMAIN 124 223
FT DISULFID 126 221
FT DISULFID 198 213
FT DISULFID 223 AA; 24255 MW; 7D66D946DCDA00 CRC64;
SQ SEQUENCE 223 AA; 24255 MW; 7D66D946DCDA00 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLV 1160
DB 5 LLLALLV 11

RESULT 71
EMBP_RAT ID EMBL RAT STANDARD; PRT; 227 AA.
AC Q63189;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eosinophil granule major basic protein precursor (MBP).
GN PRG2 OR MBP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Bone marrow;
RX MEDLINE=56138543; PubMed=8547309;
RA Nittoh T., Watanabe M., Okayama H., Misawa S., Isobe Y., Hayashi H.,
RA Mue S., Ohuchi K.;
RT "Cloning of cDNA for rat eosinophil major basic protein.";
RL Biochim. Biophys. Acta 1264:261-264(1995).
CC -!- FUNCTION: Cytotoxic and helminthotoxin. MBP also induces
CC noncytolytic histamine release from basophils. It is involved in
CC antiparasitic defense mechanisms and immune hypersensitivity
CC reactions (By similarity).
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
CC granule (crystalloid core) (By similarity).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D50568; BAA09129.1; -.
DR HSP; P13727; IHB0.
DR InterPro; IPR002352; Emaior_basic.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR PRINTS; PR00770; EMAJORBASIC.

```

```

DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
DR PROSITE; PS50041; C-TYPE LECTIN 2; 1.
KW Eosinophil; Signal; Immune response; Antibiotic; Lectin.
FT SIGNAL 1 16
FT PROPEP 17 110
FT CHAIN 111 227
FT DOMAIN 128 227
FT DISULFID 130 225
FT DISULFID 202 217
FT DOMAIN 59 68
FT DISULFID 227 AA; 25129 MW; 342B8514090DE82 CRC64;
SQ SEQUENCE 227 AA; 25129 MW; 342B8514090DE82 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 227;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLV 1160
DB 5 LLLALLV 11

RESULT 72
EMBL_CAVPO ID EMBL_CAVPO STANDARD; PRT; 233 AA.
AC P22032;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eosinophil granule major basic protein 1 precursor (MBP-1).
GN MBP1.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 115-162.
RC TISSUE=Eosinophil;
RX MEDLINE=91160746; PubMed=1705901;
RA Aoki I., Shindoh Y., Nishida T., Nakai S., Hong Y.-M., Mio M.,
RA Saito T., Tasaka K.;
RT "Sequencing and cloning of the cDNA of guinea pig eosinophil major
RT basic protein.";
RL FEBS Lett. 279:330-334(1991).
RN [2]
RP PARTIAL SEQUENCE.
RX MEDLINE=94092714; PubMed=8268206;
RA Hashimoto Y., Nagaoka I., Yamashita T.;
RT "Purification of the antibacterial fragments of guinea-pig major
RT basic protein.";
RL Biochim. Biophys. Acta 1203:236-242(1993).
CC -!- FUNCTION: MBP may play some important roles in the allergic
CC reactions and inflammations, since MBP is capable of releasing
CC histamine from mast cells and damaging the epithelial cells of
CC bronchial tubes. Antiparasitic and antibiotic.
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
CC granule (crystalloid core).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90251; BAA14291.1; -.
DR HSP; P13727; IHB0.
DR HSP; P13727; IHB0.
DR InterPro; IPR002352; Emaior_basic.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.

```

```
DR PRINTS: PR00770; EMBL:AB000169; AAC73749.1; -
DR SMART: SM00034; CUECT; 1.
DR PROSITE: PS00615; C-TYPE LECTIN 1; 1.
DR PROSITE: PS00641; C-TYPE LECTIN 2; 1.
KW Eosinophil; Signal; Immune response; Antibiotic; Lectin;
KW Multigene family.
FT SIGNAL 1 15 POTENTIAL.
FT PROPEP 16 114 ACIDIC.
FT CHAIN 115 233 EOSINOPHIL GRANULE MAJOR BASIC PROTEIN 1.
FT DOMAIN 132 233 C-TYPE LECTIN (SHORT FORM).
FT DISULFID 134 231 BY SIMILARITY.
FT DISULFID 208 223 BY SIMILARITY.
SQ SEQUENCE 233 AA; 26268 MW; C8D5E96D927C56C8 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALL 1159
DB 4 LLLALL 10

RESULT 73
YBEO_ECOLI STANDARD; PRT; 235 AA.
AC P77427;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ybeu.
GN YBEO OR B0648.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,
RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map."
RL DNA Res. 3:1137-155(1996).
CC -!- SIMILARITY: STRONG, TO E.COLI YBER.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; U28377; AAA69134.1; -
DR PIR; S65082; E65082.
DR EcoGene; EG12987; yqga.
DR InterPro; IPR007563; DUF554.
DR Pfam; PF04474; DUF554; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 2 22 POTENTIAL.
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 56 76 POTENTIAL.
FT TRANSMEM 102 122 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 178 198 POTENTIAL.
FT TRANSMEM 210 230 POTENTIAL.
SQ SEQUENCE 235 AA; 24614 MW; CC89D2A93FED29EA CRC64;

Query Match 0.6%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

or send an email to license@isb-sib.ch).
CC
DR EMBL; AB000169; AAC73749.1; -
DR EMBL; U28598; AAB40849.1; -
DR EMBL; D90704; BAA35295.1; -
DR EMBL; D90705; BAA35300.1; -
DR PIR; F64799; F64799.
DR EcoGene; EG13651; ybeu.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 235 AA; 27022 MW; 9A1CAFESD6A0ECF1 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 926 ELAAGSD 932
DB 202 ELAAGSD 208

RESULT 74
YQGA_ECOLI STANDARD; PRT; 235 AA.
ID YQGA_ECOLI
AC Q46831;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yqga.
GN YQGA OR B2966.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; U28377; AAA69134.1; -
DR EMBL; AE000379; AAC76003.1; -
DR PIR; S65082; E65082.
DR EcoGene; EG12987; yqga.
DR InterPro; IPR007563; DUF554.
DR Pfam; PF04474; DUF554; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 2 22 POTENTIAL.
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 56 76 POTENTIAL.
FT TRANSMEM 102 122 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 178 198 POTENTIAL.
FT TRANSMEM 210 230 POTENTIAL.
SQ SEQUENCE 235 AA; 24614 MW; CC89D2A93FED29EA CRC64;

Query Match 0.6%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1151 GGLLLA 1157  
Db 194 GGLLLA 200

RESULT 75

IR2\_MAIZE STANDARD; PRT; 236 AA.  
AC P16037;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Chloroplast 30S ribosomal protein S2.  
GN RPS2.  
OS Zea mays (Maize).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC PACAD clade; Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90175001; PubMed=2308853;  
RA Igloi G.L., Meinke A., Doery I., Koessel H.;  
RT "Nucleotide and derived amino acid sequence of rps2 from maize  
RT chloroplasts.";  
RL Nucleic Acids Res. 18:663-663(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=CV, FR9CMSR37; TISSUE=leaf;  
RX MEDLINE=90272437; PubMed=2140886;  
RA Stahl D., Rodermeier S., Subramanian A.R., Bogorad L.;  
RT "Nucleotide sequence of a 3.46 kb region of maize chloroplast DNA  
RT containing the gene cluster rpoC2-rps2-atpI-atpH.";  
RL Nucleic Acids Res. 18:3073-3074(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95395841; PubMed=7666415;  
RA Maier R.M., Neckermann K., Igloi G.L., Koessel H.;  
RT "Complete sequence of the maize chloroplast genome: gene content,  
RT hotspots of divergence and fine tuning of genetic information by  
RT transcript editing.";  
RL J. Mol. Biol. 251:614-628(1995).  
RN [4]  
RP SEQUENCE OF 1-36 FROM N.A.  
RX MEDLINE=90340289; PubMed=2381419;  
RA Igloi G.L., Meinke A., Doery I., Koessel H.;  
RT "Nucleotide sequence of the maize chloroplast rpo B/C1/C2 operon:  
RT comparison between the derived protein primary structures from  
RT various organisms with respect to functional domains.";  
RL Mol. Gen. Genet. 221:379-394(1990).  
CC -!- SIMILARITY: Belongs to the S2P family of ribosomal proteins.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X17318; CAA35198.1; -  
DR EMBL; X52270; CAA36512.1; -  
DR EMBL; X86563; CAA60279.1; -  
DR FIR; S08249; B32M2.  
DR Gramene; P16037; -  
DR MaizeDB; 66014; -  
DR HAMAP; MF\_00291; -; 1.  
DR InterPro; IPR001865; Ribosomal S2.  
DR InterPro; IPR005706; Ribosomal S2\_b/o.  
DR Pfam; PF00318; Ribosomal S2; 1.  
DR PRINTS; PR00395; RIBOSOMALS2.  
DR TIGRfam; TIGR01011; rpsB\_bact; 1.

DR PROSITE; PS00962; RIBOSOMAL\_S2\_1; 1.  
DR PROSITE; PS00963; RIBOSOMAL\_S2\_2; 1.  
KW Ribosomal protein; Chloroplast.  
SQ SEQUENCE 236 AA; 26881 MW; 80A3F99A8955FAD5 CRC64;  
Query Match 0.6%; Score 7; DB 1; Length 236;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLD 170  
Db 163 DIVIVLD 169

Search completed: June 24, 2004, 18:04:24  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 18:01:31 ; Search time 57 Seconds  
(without alignments)  
6576.064 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 1188  
Sequence: 1 MDLPRLVVALSLWPGFT.....FRSARRRRRPGLDTPKVLK 1189

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

SPREMBL\_25.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phase.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriap.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID     | Description         |
|------------|-------|---------------|--------|-----------|---------------------|
| 1          | 662   | 55.7          | 823    | 4 Q8WY18  | Q8WY18 homo sapien  |
| 2          | 122   | 10.3          | 1188   | 11 Q7TQC3 | Q7TQC3 mus musculus |
| 3          | 45    | 3.8           | 823    | 11 Q8CE84 | Q8CE84 mus musculus |
| 4          | 15    | 1.3           | 1171   | 13 Q42094 | Q42094 gallus gall  |
| 5          | 11    | 0.9           | 1288   | 11 Q8BM12 | Q8BM12 mus musculus |
| 6          | 10    | 0.8           | 116    | 13 Q31670 | Q31670 xenopus lae  |
| 7          | 10    | 0.8           | 399    | 16 Q8XSH5 | Q8XSH5 raietonia s  |
| 8          | 9     | 0.8           | 164    | 4 Q8N112  | Q8N112 homo sapien  |
| 9          | 9     | 0.8           | 315    | 16 Q8YCR3 | Q8YCR3 bruceella me |
| 10         | 9     | 0.8           | 316    | 16 Q8FVL4 | Q8FVL4 bruceella su |
| 11         | 9     | 0.8           | 400    | 6 Q8HY16  | Q8HY16 cebus apell  |
| 12         | 9     | 0.8           | 413    | 16 Q89L58 | Q89L58 bradyrhizob  |
| 13         | 9     | 0.8           | 457    | 13 Q73804 | Q73804 fugu rubrip  |
| 14         | 9     | 0.8           | 607    | 10 Q39775 | Q39775 gnetum gnet  |
| 15         | 9     | 0.8           | 780    | 13 Q06271 | Q06271 xenopus lae  |
| 16         | 8     | 0.7           | 54     | 6 Q9TS65  | Q9TS65 canis famil  |

|    |   |     |     |           |                     |
|----|---|-----|-----|-----------|---------------------|
| 17 | 8 | 0.7 | 58  | 16 Q88NW3 | Q88NW3 pseudomonas  |
| 18 | 8 | 0.7 | 76  | 11 Q8CRU3 | Q8CRU3 mus musculus |
| 19 | 8 | 0.7 | 78  | 6 Q9N241  | Q9N241 saimiri bol  |
| 20 | 8 | 0.7 | 87  | 6 Q9N242  | Q9N242 ateles sp.   |
| 21 | 8 | 0.7 | 88  | 4 Q9587   | Q9587 homo sapien   |
| 22 | 8 | 0.7 | 91  | 6 Q9N244  | Q9N244 nasalis lar  |
| 23 | 8 | 0.7 | 91  | 6 Q9N245  | Q9N245 macaca mula  |
| 24 | 8 | 0.7 | 91  | 6 Q97517  | Q97517 gorilla gor  |
| 25 | 8 | 0.7 | 91  | 6 Q97515  | Q97515 pan paniscu  |
| 26 | 8 | 0.7 | 91  | 6 Q97521  | Q97521 pongo pygma  |
| 27 | 8 | 0.7 | 91  | 6 Q9N243  | Q9N243 presbytis c  |
| 28 | 8 | 0.7 | 91  | 6 Q9N246  | Q9N246 papio hamad  |
| 29 | 8 | 0.7 | 91  | 6 Q9N247  | Q9N247 hylobates s  |
| 30 | 8 | 0.7 | 95  | 4 Q9586   | Q9586 homo sapien   |
| 31 | 8 | 0.7 | 95  | 6 Q97514  | Q97514 pan troglod  |
| 32 | 8 | 0.7 | 95  | 6 Q9N249  | Q9N249 hylobates k  |
| 33 | 8 | 0.7 | 95  | 6 Q9N248  | Q9N248 hylobates l  |
| 34 | 8 | 0.7 | 95  | 6 Q97516  | Q97516 gorilla gor  |
| 35 | 8 | 0.7 | 95  | 6 Q97520  | Q97520 pongo pygma  |
| 36 | 8 | 0.7 | 95  | 6 Q97519  | Q97519 pongo pygma  |
| 37 | 8 | 0.7 | 99  | 6 Q97518  | Q97518 pongo pygma  |
| 38 | 8 | 0.7 | 104 | 16 Q8E9F1 | Q8E9F1 shewanella   |
| 39 | 8 | 0.7 | 106 | 10 Q9SDS4 | Q9SDS4 capsicum an  |
| 40 | 8 | 0.7 | 112 | 11 Q8RIE9 | Q8RIE9 mus musculu  |
| 41 | 8 | 0.7 | 112 | 11 Q8J2X1 | Q8J2X1 mus musculu  |
| 42 | 8 | 0.7 | 112 | 11 Q7TNY5 | Q7TNY5 mus musculu  |
| 43 | 8 | 0.7 | 112 | 11 Q7TNN0 | Q7TNN0 mus musculu  |
| 44 | 8 | 0.7 | 112 | 11 Q7TNN7 | Q7TNN7 mus musculu  |
| 45 | 8 | 0.7 | 114 | 10 Q9ZF19 | Q9ZF19 capsicum an  |
| 46 | 8 | 0.7 | 127 | 16 Q9RSB5 | Q9RSB5 deinococcus  |
| 47 | 8 | 0.7 | 127 | 16 Q82U09 | Q82U09 nitrosomona  |
| 48 | 8 | 0.7 | 137 | 11 Q8C442 | Q8C442 mus musculu  |
| 49 | 8 | 0.7 | 141 | 11 Q61692 | Q61692 mus musculu  |
| 50 | 8 | 0.7 | 144 | 11 Q7TIM3 | Q7TIM3 bothrops ja  |
| 51 | 8 | 0.7 | 146 | 13 Q8Q590 | Q8Q590 bothrops in  |
| 52 | 8 | 0.7 | 161 | 10 Q7XR40 | Q7XR40 oryza sativ  |
| 53 | 8 | 0.7 | 174 | 15 Q98MW8 | Q98MW8 rhizobium l  |
| 54 | 8 | 0.7 | 181 | 13 F9786  | F9786 gallus gall   |
| 55 | 8 | 0.7 | 181 | 13 Q90Y12 | Q90Y12 crotalus du  |
| 56 | 8 | 0.7 | 181 | 13 Q90Y11 | Q90Y11 crotalus du  |
| 57 | 8 | 0.7 | 222 | 17 Q8U4N4 | Q8U4N4 pyrococcus   |
| 58 | 8 | 0.7 | 224 | 15 Q7V514 | Q7V514 prochloroco  |
| 59 | 8 | 0.7 | 224 | 17 Q57835 | Q57835 pyrococcus   |
| 60 | 8 | 0.7 | 226 | 16 F74604 | F74604 synechocyst  |
| 61 | 8 | 0.7 | 227 | 16 Q882P2 | Q882P2 pseudomonas  |
| 62 | 8 | 0.7 | 231 | 16 Q9AJX4 | Q9AJX4 streptomyce  |
| 63 | 8 | 0.7 | 235 | 16 Q98GF8 | Q98GF8 rhizobium l  |
| 64 | 8 | 0.7 | 245 | 5 Q20220  | Q20220 caenorhabdi  |
| 65 | 8 | 0.7 | 257 | 16 Q8Y1R9 | Q8Y1R9 raietonia s  |
| 66 | 8 | 0.7 | 265 | 13 Q8Q391 | Q8Q391 bothrops in  |
| 67 | 8 | 0.7 | 265 | 13 Q9PW56 | Q9PW56 bothrops ja  |
| 68 | 8 | 0.7 | 272 | 7 Q861M5  | Q861M5 equus cabal  |
| 69 | 8 | 0.7 | 272 | 7 Q861K3  | Q861K3 equus cabal  |
| 70 | 8 | 0.7 | 278 | 16 Q8P8N6 | Q8P8N6 xanthomonas  |
| 71 | 8 | 0.7 | 290 | 17 Q8TPV4 | Q8TPV4 methanopyru  |
| 72 | 8 | 0.7 | 299 | 10 Q94K03 | Q94K03 arabidopsis  |
| 73 | 8 | 0.7 | 307 | 16 Q98BD8 | Q98BD8 rhizobium l  |
| 74 | 8 | 0.7 | 308 | 16 Q92S85 | Q92S85 rhizobium m  |
| 75 | 8 | 0.7 | 323 | 11 Q8CB84 | Q8CB84 mus musculu  |
| 76 | 8 | 0.7 | 328 | 16 Q82YY0 | Q82YY0 enterococcu  |
| 77 | 8 | 0.7 | 339 | 2 Q9KWB6  | Q9KWB6 agrobacteri  |
| 78 | 8 | 0.7 | 348 | 4 Q8TES5  | Q8TES5 homo sapien  |
| 79 | 8 | 0.7 | 350 | 4 Q7Z7Q5  | Q7Z7Q5 homo sapien  |
| 80 | 8 | 0.7 | 356 | 16 Q8U102 | Q8U102 agrobacteri  |
| 81 | 8 | 0.7 | 360 | 10 Q9M146 | Q9M146 arabidopsis  |
| 82 | 8 | 0.7 | 371 | 16 Q89W76 | Q89W76 bradyrhizob  |
| 83 | 8 | 0.7 | 392 | 16 Q881M1 | Q881M1 pseudomonas  |
| 84 | 8 | 0.7 | 405 | 10 Q94FT6 | Q94FT6 fragaria an  |
| 85 | 8 | 0.7 | 405 | 10 Q24416 | Q24416 fragaria an  |
| 86 | 8 | 0.7 | 415 | 16 Q915T3 | Q915T3 pseudomonas  |
| 87 | 8 | 0.7 | 415 | 16 Q8Q8S5 | Q8Q8S5 pseudomonas  |
| 88 | 8 | 0.7 | 415 | 16 Q88A33 | Q88A33 pseudomonas  |
| 89 | 8 | 0.7 | 419 | 4 Q8NGM5  | Q8NGM5 homo sapien  |

90 Q8P60 xanthomonas  
91 Q9P3U1 schizosacch  
92 Q8ye57 brucella me  
93 Q8xys8 brucella su  
94 Q04611 arabidopsis  
95 Q8ewh6 mycoplasma  
96 Q9eyg0 amycolatops  
97 Q8ziz0 streptomyc  
98 Q52642 pseudomonas  
99 Q51991 pseudomonas  
100 Q64584 rattus norv

ALIGNMENTS

RESULT 1

Q8WY18 PRELIMINARY; PRT; 823 AA.  
AC Q8WY18  
DT 01-MAR-2002 (Tremblrel. 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE XSTP018.  
GN MST018.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Aorta;  
RA Xu Y.Y., Sun L.Z., Wu Q.Y., Liu Y.Q., Liu B., Zhao B., Wang X.Y.,  
Song L., Ye J., Sheng H., Gao Y., Zhang C.L., Zhang J., Wei Y.J.,  
Sun Y.H., Jiang Y.X., Zhao X.W., Liu S., Liu L.S., Ding J.P.,  
Gao R.L., Qiang B.O., Yuan J.G., Liew C.C., Zhao M.S., Hui R.T.;  
Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
RA EMBL: AF111799; AAC139001.1;  
DR GO: 0008305; C: integrin complex; IEA.  
DR GO: 0004895; F: cell-adhesion receptor activity; IEA.  
DR GO: 0007160; P: cell-matrix adhesion; IEA.  
DR InterPro: IPR000413; Integrin\_alpha.  
DR Pfam: PF01839; FG-GAP; 3.  
DR PRINTS: PR01185; INTEGRIN.  
DR SMART: SM00191; Int\_alpha; 4.  
SQ SEQUENCE 823 AA; 92672 MW; DE4E78079DCD4925 CRC64;

Query Match 55.7%; Score 662; DB 4; Length 823;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 762; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
366 MSOTGFSHVVDGVLGAVGAYDNGAVLKTSAGKVIPLRESYLKPEELKNGAYL 425  
1 MSOTGFSHVVDGVLGAVGAYDNGAVLKTSAGKVIPLRESYLKPEELKNGAYL 60  
426 GYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMNNRSLTIHQAMRGQIGSYFGSEIT 485  
61 GYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMNNRSLTIHQAMRGQIGSYFGSEIT 120  
486 SVDIDGCVDTLLVGAAPKVFNEGRGKVVYELRONRVYNGTLKDSHYONARFGSS 545  
121 SVDIDGCVDTLLVGAAPKVFNEGRGKVVYELRONRVYNGTLKDSHYONARFGSS 180  
546 IASVRDLNQDSYNDVVVVGAPLEDNHAGAIYFHGFRGSILKTPKQRTASELATGLQYFG 605  
181 IASVRDLNQDSYNDVVVVGAPLEDNHAGAIYFHGFRGSILKTPKQRTASELATGLQYFG 240  
606 CSIHGQDLNEDGLIDLAVAGLGNVILWSRPVQVQINASLHFPSPKINIFHRCKSGRD 665  
241 CSIHGQDLNEDGLIDLAVAGLGNVILWSRPVQVQINASLHFPSPKINIFHRCKSGRD 300  
566 ATCLAAFLCPTPIFLAPHFOTTTGIRYNATMDERRYTPRAHLDEGGDRFTNRVLLSSG 725

301 ATCLAAFLCPTPIFLAPHFOTTTGIRYNATMDERRYTPRAHLDEGGDRFTNRVLLSSG 360  
726 QELCERINPHVLDADYVVKPVTFESVEYSLEDPDHGPMDDGMPPTLTRVSVFPWNGCNEDE 785  
361 QELCERINPHVLDADYVVKPVTFESVEYSLEDPDHGPMDDGMPPTLTRVSVFPWNGCNEDE 420  
786 HCVPLDLDAKSDLPAMBYCORVLRKPAQDCSAYTLSFDTTVFIISSTRQVAVEATLE 845  
421 HCVPLDLDAKSDLPAMBYCORVLRKPAQDCSAYTLSFDTTVFIISSTRQVAVEATLE 480  
846 NRGENAYTVLNISOSANLOFASLLOKEDSGSTECVNEERLQKQVNSVYPPFRACK 905  
481 NRGENAYTVLNISOSANLOFASLLOKEDSGSTECVNEERLQKQVNSVYPPFRACK 540  
906 VAFRLDSEFSKSIPLHLEIEIAGAAGSDNERDSTKEDNVAPLRPHLKYEADVLFTRSSSL 965  
541 VAFRLDSEFSKSIPLHLEIEIAGAAGSDNERDSTKEDNVAPLRPHLKYEADVLFTRSSSL 600  
966 SHYEVKLASSLERYDGIQPPFSCIFRIQNLGLPIHGMMKTIPIATRSNELLKLRDP 1025  
601 SHYEVKLASSLERYDGIQPPFSCIFRIQNLGLPIHGMMKTIPIATRSNELLKLRDP 660  
1026 LTDEANTSCNIWGNSTERYPTPVEDLRRAPQLNHSNDVVSINIRLVPNOEINPHLL 1085  
661 LTDEANTSCNIWGNSTERYPTPVEDLRRAPQLNHSNDVVSINIRLVPNOEINPHLL 720  
1086 GNLWLRSLKALKYKSMKIMVNAALQRFHSPPIFREDDPSRQI 1128  
721 GNLWLRSLKALKYKSMKIMVNAALQRFHSPPIFREDDPSRQI 763

RESULT 2

Q7TQC3 PRELIMINARY; PRT; 1188 AA.  
AC Q7TQC3  
DT 01-OCT-2003 (Tremblrel. 25, Created)  
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE All integrin.  
GN ITGAL1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Tiger C.-F., Popova S.N., Velling T., Ekman S., Forsberg E.,  
Gullberg D.;  
RT "allb1 integrin is important for mesenchymal cell function:  
elimination of allb1 leads to dwarfism."  
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Johansson M., Popova S.N.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY124460; AAM62130.1; -;  
KW Integrin.  
SQ SEQUENCE 1188 AA; 133067 MW; 69B2BF20D75E4271 CRC64;

Query Match 10.3%; Score 122; DB 11; Length 1188;  
Best Local Similarity 100.0%; Pred. No. 8.3e-118;  
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
81 GNCTKLNLRVTLNSVSRKONMRLGLSLATNPKNDFLACPLWSHCGSSYTTGMS 140  
81 GNCTKLNLRVTLNSVSRKONMRLGLSLATNPKNDFLACPLWSHCGSSYTTGMS 140  
141 RVNSFRSKTVAPALQRCOTMDIVIVLDGNSIYPWVEVQHLNLIKXFIIGPQIQ 200  
141 RVNSFRSKTVAPALQRCOTMDIVIVLDGNSIYPWVEVQHLNLIKXFIIGPQIQ 200  
201 VG 202  
||

Db 201 VG 202

RESULT 3

Q8CEB4 PRELIMINARY; PRT; 823 AA.

AC Q8CEB4;

DT 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE MSP018 homolog.

GN 4732459H24RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

EN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CS7BL/6J; TISSUE=Skin;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium.

RA "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs."

RL Nature 420:563-573 (2002).

DR EMBL; AK028821; BAC26137.1; --

DR MGD; MGI:2442114; 4732459H24RIK.

DR GO; GO:0008305; C:integrin complex; IEA.

DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.

DR GO; GO:0007160; P:cell-matrix adhesion; IEA.

DR InterPro; IPR000413; Integrin\_alpha.

DR Pfam; PF01839; FG-GAP; 3.

DR PRINTS; PR01185; INTEGRINA.

DR SMART; SM00191; Int. alpha; 4.

DR SMART; SM00191; Int. alpha; 4.

DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.

DR PROSITE; PS00234; VWFA; 1.

DR Integrin.

SW SEQUENCE 1171 AA; 130228 MW; B505A4D65F09736E CRC64;

Qy 164 DIVVLGDSNSIYFW 178

Db 162 DIVVLGDSNSIYFW 176

RESULT 5

Q8BM12 PRELIMINARY; PRT; 288 AA.

AC Q8BM12

DT 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Integrin alpha-10 precursor homolog (Fragment).

GN ITGA10.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

EN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CS7BL/6J; TISSUE=Skin;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium.

RA "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs."

RL Nature 420:563-573 (2002).

DR EMBL; AK037222; BAC29761.1; --

DR MGD; MGI:2153482; Itga10.

DR GO; GO:0008305; C:integrin complex; IEA.

DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.

DR GO; GO:0007160; P:cell-matrix adhesion; IEA.

DR InterPro; IPR000413; Integrin\_alpha.

DR Pfam; PF00357; Integrin\_A; 1.

DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.

FT NON TER 1

SW SEQUENCE 288 AA; 31949 MW; 3F6200F9C9475BE5 CRC64;

Qy 1150 LGGLLLALLV 1160

Db 252 LGGLLLALLV 262

RESULT 6

Q91670 PRELIMINARY; PRT; 116 AA.

AC Q91670;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

Qy 1150 LGGLLLALLV 1160

Db 252 LGGLLLALLV 262

Db 201 VG 202

RESULT 4

Q42094 PRELIMINARY; PRT; 1171 AA.

AC Q42094;

DT 01-JAN-1998 (TREMBlrel. 05, Created)

DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE ALP11 integrin.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI\_TaxID=9031;

EN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Gizzard;

RX MEDLINE=97476270; PubMed=9334246;

RA Obata H.; Hayashi K.; Nishida W.; Momiyama T.; Uchida A.; Ochi T.;

RA Sobue K.;

RT "Smooth muscle cell phenotype-dependent transcriptional regulation of

RT the alp11 integrin gene."

RL J. Biol. Chem. 272:26643-26651 (1997).

DR EMBL; AB000470; BAA23160.1; --

DR EMBL; AB000471; BAA23161.1; --

DR PIR; A55348; A55348.

DR HSSP; P17301; 1A0X.

DR GO; GO:0008305; C:integrin complex; IEA.

DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.

Qy 766 GPTTILRVSPWNCNDEHCVPLVLDARSDLPTAMEYQCVL 810

Db 401 GPTTILRVSPWNCNDEHCVPLVLDARSDLPTAMEYQCVL 445

Query Match 3.8%; Score 45; DB 11; Length 823;

Best Local Similarity 100.0%; Pred. No. 2.5e-37; Indels 0; Gaps 0;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 GPTTILRVSPWNCNDEHCVPLVLDARSDLPTAMEYQCVL 810

Db 401 GPTTILRVSPWNCNDEHCVPLVLDARSDLPTAMEYQCVL 445

RESULT 4

Q42094 PRELIMINARY; PRT; 1171 AA.

AC Q42094;

DT 01-JAN-1998 (TREMBlrel. 05, Created)

DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE ALP11 integrin.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI\_TaxID=9031;

EN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Gizzard;

RX MEDLINE=97476270; PubMed=9334246;

RA Obata H.; Hayashi K.; Nishida W.; Momiyama T.; Uchida A.; Ochi T.;

RA Sobue K.;

RT "Smooth muscle cell phenotype-dependent transcriptional regulation of

RT the alp11 integrin gene."

RL J. Biol. Chem. 272:26643-26651 (1997).

DR EMBL; AB000470; BAA23160.1; --

DR EMBL; AB000471; BAA23161.1; --

DR PIR; A55348; A55348.

DR HSSP; P17301; 1A0X.

DR GO; GO:0008305; C:integrin complex; IEA.

DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.

```

01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Integrin alpha-1 (Fragment).
GN GENE 4.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312514; PubMed=8700860;
RA Brown D.D., Wang Z., Furlow J.D., Kanamori A., Schwartzman R.A.,
RA Remo B.F., Binder A.;
RT "The thyroid hormone-induced tail resorption program during Xenopus
RT laevis metamorphosis.";
RL EMBL; U44025; AAC59878.1; -.
DR GO; GO:0008305; C:Integrin complex; IEA.
DR GO; GO:0004895; F:Cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:Cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR SMART; SM00191; Int_alpha; 1.
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 12749 MW; 2A8A5CA3E83DD07E CRC64;

Query Match 0.8%; Score 10; DB 13; Length 116;
Best Local Similarity 100.0%; Pred. No. 0.2; Mismatches 0; Gaps 0;
Matches 10; Conservative 0; Indels 0;

Qy 39 AFFGYTQQH 48
Db 39 AFFGYTQQH 48
|||||
|

RESULT 7
Q8XSH5 PRELIMINARY; PRT; 399 AA.
AC Q8XSH5;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative transport transmembrane protein.
GN R5F0499 OR R500378.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OG Plasmid megaplasmid.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Axlatt M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguler P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646079; CAD17650.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0006457; P:protein folding; IEA.
DR InterPro; IPR001175; FKBP_PPase.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS00453; FKBP_PPase_1; 1.
DR PROSITE; PS50850; MFS; 1.
DR Plasmid; Complete proteome.
SQ SEQUENCE 399 AA; 42148 MW; 096BB97E83DCA7D1 CRC64;

Query Match 0.8%; Score 10; DB 16; Length 399;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1152 GLLLLALLVL 1161
Db 161 GLLLLALLVL 170
|||||
|

RESULT 8
Q8N112 PRELIMINARY; PRT; 164 AA.
AC Q8N112;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-WAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical gene LOC132228 (Hypothetical protein FLJ38608).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX Tissue=Brain;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC Tissue=Heart;
RA Nishi T., Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,
RA Sugiyama T., Irie K., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEO human cDNA sequencing project";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028000; AAH28000.1; -.
DR EMBL; AK095927; BAC04652.1; -.
KW Hypothetical protein.
SQ SEQUENCE 164 AA; 17868 MW; 809993458714252E CRC64;

Query Match 0.8%; Score 9; DB 4; Length 164;
Best Local Similarity 100.0%; Pred. No. 3; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

Qy 1153 LLLALLVL 1161
Db 99 LLLALLVL 107
|||||
|

RESULT 9
Q8YCR3 PRELIMINARY; PRT; 315 AA.
AC Q8YCR3;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE PERMEASE.
GN EME110465.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Muej C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonaki L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Ietesson J.-J.,
RA Haseikorn R., Kyrpides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";

```



RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).  
DR EMBL; AB009603; AAL53707.1; -.  
DR FIR; AH3567; AH3567.  
DR InterPro; IPR002781; DUF81.  
DR Pfam; PF01925; DUF81; 1.  
KW Complete proteome.  
SQ SEQUENCE 315 AA; 33486 MW; 7A34836134FB2F7D CRC64;  
  
Query Match 0.8%; Score 9; DB 16; Length 315;  
Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1154 LLLALLVLA 1162  
Db 283 LLLALLVLA 291  
|||||  
|||  
  
RESULT 10  
Q8FVL4 PRELIMINARY; PRT; 316 AA.  
AC Q8FVL4;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Membrane protein, putative.  
GN BRA0823.  
OS Brucella suis.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Brucellaceae; Brucella.  
OX NCBI\_TaxID=29461;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1330 / Biovar 1;  
RX MEDLINE=22247741; PubMed=12271122;  
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,  
RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,  
RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R., S.E.,  
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,  
RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,  
RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;  
RT "The Brucella suis genome reveals fundamental similarities between  
RT animal and plant pathogens and symbionts."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).  
DR EMBL; AB014576; AAN3998.1; -.  
DR TIGR; BRA0823; -.  
DR InterPro; IPR002781; DUF81.  
DR Pfam; PF01925; DUF81; 1.  
KW Complete proteome.  
SQ SEQUENCE 316 AA; 33671 MW; 4DC0DC32D8AB9807 CRC64;  
  
Query Match 0.8%; Score 9; DB 16; Length 316;  
Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1154 LLLALLVLA 1162  
Db 284 LLLALLVLA 292  
|||||  
|||  
  
RESULT 11  
Q8HY16 PRELIMINARY; PRT; 400 AA.  
AC Q8HY16;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Poliovirus receptor.  
GN PVR.  
OS Cebus apella (Brown-capped capuchin).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebus.  
OX NCBI\_TaxID=9515;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=USDA 110;  
RX MEDLINE=22484998; PubMed=12597275;  
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,  
RA Sasamoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,  
RA Kohara M., Matsumoto M., Shimpo S., Teurouka H., Wada T., Yamada M.,  
RA Tabata S.;  
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
RT Bradyrhizobium japonicum USDA110";  
RL DNA Res. 9:189-197(2002).  
DR EMBL; AP005951; BAC49730.1; -.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0016811; F:hydrolase activity, acting on carbon-nitrog. .; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR004304; FmdA\_AmdA.  
DR Pfam; PF03069; FmdA\_AmdA; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
KW Complete proteome.  
SQ SEQUENCE 413 AA; 45159 MW; E2F309DCC2E8D82B CRC64;  
  
Query Match 0.8%; Score 9; DB 16; Length 413;  
Best Local Similarity 100.0%; Pred. No. 6.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
RP SEQUENCE FROM N.A.  
RA Ida-Hosonuma M., Sasaki Y., Toyoda H., Nomoto A., Gotoh O.,  
RA Yonekawa H., Koike S.;  
RT "Host range of poliovirus is restricted to simians because of a rapid  
RT sequence change of the poliovirus receptor gene during evolution."  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; AB086131; BAC41711.1; -.  
DR EMBL; AB086124; BAC41711.1; JOINED.  
DR EMBL; AB086125; BAC41711.1; JOINED.  
DR EMBL; AB086126; BAC41711.1; JOINED.  
DR EMBL; AB086127; BAC41711.1; JOINED.  
DR EMBL; AB086128; BAC41711.1; JOINED.  
DR EMBL; AB086129; BAC41711.1; JOINED.  
DR EMBL; AB086130; BAC41711.1; JOINED.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR InterPro; IPR001064; Crystallin.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 3.  
DR SMART; SMO0406; IGV; 1.  
DR PROSITE; PS00225; CRYSTALLIN\_BETAGAMMA; 1.  
DR PROSITE; PS00835; IG\_LIKE; 3.  
KW Receptor.  
SQ SEQUENCE 400 AA; 43322 MW; C6C3A59A0A285F3A CRC64;  
  
Query Match 0.8%; Score 9; DB 6; Length 400;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1153 LLLALLVLA 1161  
Db 9 LLLALLVLA 17  
|||||  
|||  
  
RESULT 12  
Q89LS8 PRELIMINARY; PRT; 413 AA.  
AC Q89LS8;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE B1r4465 protein.  
GN B1r4465.  
OS Bradyrhizobium japonicum.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Bradyrhizobium.  
OX NCBI\_TaxID=375;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=USDA 110;  
RX MEDLINE=22484998; PubMed=12597275;  
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,  
RA Sasamoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,  
RA Kohara M., Matsumoto M., Shimpo S., Teurouka H., Wada T., Yamada M.,  
RA Tabata S.;  
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
RT Bradyrhizobium japonicum USDA110";  
RL DNA Res. 9:189-197(2002).  
DR EMBL; AP005951; BAC49730.1; -.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0016811; F:hydrolase activity, acting on carbon-nitrog. .; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR004304; FmdA\_AmdA.  
DR Pfam; PF03069; FmdA\_AmdA; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
KW Complete proteome.  
SQ SEQUENCE 413 AA; 45159 MW; E2F309DCC2E8D82B CRC64;

QY 148 FSKTVAPAL 156  
Db 18 FSKTVAPAL 26

RESULT 13

O73804 O73804 PRELIMINARY; PRT; 497 AA.  
AC O73804; 1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
OC Tetraodontidae; Tetraodontidae; Takifugu.  
OX NCBI\_TaxID=31033;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=99177347; PubMed=10077531;  
RA Gellner K., Brenner S.;  
RT "Analysis of 143 kb of genomic DNA around the wnt1 locus of Fugu  
rubripes.";  
RL Genome Res. 9:251-258(1999).  
DR EMBL; AF056116; AAC34385.1;  
DR InterPro; IPR008075; Lipocalin1\_recep.  
DR InterPro; IPR006876; LMBR1.  
DR Pfam; PF04791; LMBR1; 1.  
DR PRINTS; PR01692; LIPOCALINMR.  
KW Hypothetical protein.  
SQ SEQUENCE 497 AA; 56376 MW; 802743E398383B64 CRC64;

Query Match 0.8%; Score 9; DB 13; Length 497;  
Best Local Similarity 100.0%; Pred. No. 8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVL 1161  
Db 169 LLLALLVL 177

RESULT 14

O39775 O39775 PRELIMINARY; PRT; 607 AA.  
AC O39775;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Legumin, 11S globulin.  
OS Gnetum gnemon (Bago).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Gnetales; Gnetales; Gnetales; Gnetales; Gnetales.  
OX NCBI\_TaxID=3382;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX TISSUE=Seed;  
RX MEDLINE=9840765; PubMed=9767693;  
RA Shatov A.D., Braun H., Chesnokov Y.V., Horstmann C., Kakhovskaya I.A.,  
RA Baumlain H.;  
RT "Sequence peculiarity of Gnetalean legumin-like seed storage  
proteins.";  
RL J. Mol. Evol. 47:486-492(1998).  
DR EMBL; 250779; CAA90642.1;  
DR PIR; S60658; S60658.  
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.  
DR InterPro; IPR006045; Cupin.  
DR Pfam; PF00190; Cupin; 2.  
DR PRINTS; PR00439; 11SGLOBULIN.  
SQ SEQUENCE 607 AA; 68910 MW; 7AD6DDC504FB3A60 CRC64;

Query Match 0.8%; Score 9; DB 10; Length 607;  
Best Local Similarity 100.0%; Pred. No. 9.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVL 1161  
Db 10 LLLALLVL 18

RESULT 15

O06271 O06271 PRELIMINARY; PRT; 780 AA.  
AC O06271;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Integrin alpha-2 subunit (Fragment).  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
OC Xenopodidae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RN SEQUENCE FROM N.A.  
RA Meng F., Desimone D.W.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RN SEQUENCE OF 95-168 FROM N.A.  
RX MEDLINE=94008528; PubMed=8404528;  
RA Whittaker C.A., Desimone D.W.;  
RT "Integrin alpha subunit mRNAs are differentially expressed in early  
Xenopus embryos.";  
RL Development 117:1239-1249(1993).  
DR EMBL; L43058; AAA69770.1;  
DR PIR; I51524; I51524.  
DR GO; GO:0008305; C:integrin complex; IEA.  
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.  
DR GO; GO:0007150; P:cell-matrix adhesion; IEA.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR Pfam; PF01839; FG-GAP; 3.  
DR PRINTS; PR01185; INTEGRIN\_A; 1.  
DR SMART; SM00191; Int\_alpha; 4.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
FT NON\_TER 1  
SQ SEQUENCE 780 AA; 87017 MW; 9518B18C2B6BF637 CRC64;

Query Match 0.8%; Score 9; DB 13; Length 780;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGP 503  
Db 108 TDVLLVGP 116

RESULT 16

O9TS65 O9TS65 PRELIMINARY; PRT; 54 AA.  
AC O9TS65;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE LEUKOINTEGRIN alpha D beta 2 (Fragments).  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RN SEQUENCE.  
RX MEDLINE=95325609; PubMed=7541420;  
RA Danilenko D.M., Rossitto P.V., Van der Vieren M., Le Trong H.,

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;  
CD24A. Mus musculus (Mouse).

QY 1154 LLLALLVL 1161  
|||  
Db 64 LLLALLVL 71

|    |             |   |
|----|-------------|---|
| AC | Q9N242;     |   |
| DT | 01-OCT-2000 | (TRENELrel. 15, Created)                |
| DT | 01-OCT-2000 | (TRENELrel. 15, last sequence update)   |
| DT | 01-JUN-2003 | (TRENELrel. 24, Last annotation update) |

DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Ateles sp. (Spider monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateleinae; Ateles.  
OX NCBI\_TaxID=9511;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
RT primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125669; AAF64384.1; -;  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCRHHODOPSN.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER 87 87  
SQ SEQUENCE 87 AA; 8388 MW; 64574359DAC555D4 CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 87;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 73 LLLALLVL 80  
|||||  
  
RESULT 21  
ID Q99587 PRELIMINARY; PRT; 88 AA.  
AC Q99587;  
DT 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96362071; PubMed=8746407;  
RA Cichon S., Nothen M.M., Catalano M., Di Bella D., Maier W.,  
RA Lichtermann D., Mirges J., Albus M., Borrmann M., Franzeck E., et al;  
RT "Identification of two novel polymorphisms and a rare deletion variant  
RT in the human dopamine D4 receptor gene.";  
RL Psychiatr. Genet. 5:97-103(1995).  
DR EMBL; S82918; AAB46803.1; -;  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCRHHODOPSN.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER 88 88  
SQ SEQUENCE 88 AA; 8482 MW; 314093AB4678CEDF CRC64;  
  
Query Match 0.7%; Score 8; DB 4; Length 88;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161

DB 74 LLLALLVL 81  
|||||  
  
RESULT 22  
Q9N244 PRELIMINARY; PRT; 91 AA.  
ID Q9N244;  
AC Q9N244;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Nasalis larvatus (Proboscis monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Colobinae;  
OC Nasalis.  
OX NCBI\_TaxID=43780;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
RT primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125667; AAF64382.1; -;  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCRHHODOPSN.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER 91 91  
SQ SEQUENCE 91 AA; 8827 MW; 67B53D0764BD439C CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 91;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 77 LLLALLVL 84  
|||||  
  
RESULT 23  
Q9N245 PRELIMINARY; PRT; 91 AA.  
ID Q9N245;  
AC Q9N245;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  
OC Cercopitheinae; Macaca.  
OX NCBI\_TaxID=9544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
RT primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125666; AAF64381.1; -;  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.

```

DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 91
SQ SEQUENCE 91 AA; 8728 MW; 7CF88CAC7E0CF939C CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 6; Length 91;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 77 LLLALLVL 84

RESULT 24
O97517 PRELIMINARY; PRT; 91 AA.
AC O97517;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4
OS Gorilla gorilla (Gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010297; AAC67225.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 91
SQ SEQUENCE 91 AA; 8763 MW; 765DBF6B3E385E299 CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 6; Length 91;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 77 LLLALLVL 84

RESULT 26
O97521 PRELIMINARY; PRT; 91 AA.
AC O97521;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010301; AAC67229.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 91
SQ SEQUENCE 91 AA; 8934 MW; 93CD0F1CF8E3E82C CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 6; Length 91;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 77 LLLALLVL 84

RESULT 27
O9N243 PRELIMINARY; PRT; 91 AA.
ID O9N243

```

AC Q9N243; 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Presbytis cristata (Silvered langur).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;  
OC Presbytis.  
OX NCBI\_TaxID=36232;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
RT Primates";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125668; AAF64380.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm1; 1.  
DR PRINTS; PR00237; GPCRHHODPSN.  
DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER  
SQ SEQUENCE 91 AA; 8927 MW; 67E53D0764B0439C CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 91;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 77 LLLALLVL 84  
  
RESULT 28  
ID Q9N246 PRELIMINARY; PRT; 91 AA.  
AC Q9N246;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Papio hamadryas (Hamadryas baboon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Papio.  
OX NCBI\_TaxID=9557;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
RT Primates";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125665; AAF64380.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm1; 1.  
DR PRINTS; PR00237; GPCRHHODPSN.  
DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER  
SQ SEQUENCE 91 AA; 8770 MW; 608F90C00E10839C CRC64;

QY 1154 LLLALLVL 1161  
DB 77 LLLALLVL 84  
  
RESULT 29  
ID Q9N247 PRELIMINARY; PRT; 93 AA.  
AC Q9N247;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Hylobates syndactylus (Siamang) (Symphalangus syndactylus).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.  
OX NCBI\_TaxID=9590;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
RT Primates";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125684; AAF64379.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm1; 1.  
DR PRINTS; PR00237; GPCRHHODPSN.  
DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER  
SQ SEQUENCE 93 AA; 8904 MW; 5D07B793F2B624E4 CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 93;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 79 LLLALLVL 86  
  
RESULT 30  
ID Q99586 PRELIMINARY; PRT; 95 AA.  
AC Q99586;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96362071; PubMed=8746407;  
RA Cichon S., Nothen M.M., Catalano M., Di Bella D., Maier W.,  
RA Lichtermann D., Minges J., Albus M., Borrmann M., Franzeck B., et al;  
RT "Identification of two novel polymorphisms and a rare deletion variant  
RT in the human dopamine D4 receptor gene";  
RL Psychiatr. Genet. 5:97-103(1995).  
DR EMBL; S82917; AAB46802.1; -.

DR GO; GO:0004952; F:dopamine receptor activity; NAS.  
DR GO; GO:0007212; P:dopamine receptor signaling pathway; NAS.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCR\_Rhodopsn.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER 95 95  
SQ SEQUENCE 95 AA; 9148 MW; 7163AF13B6156686 CRC64;  
  
Query Match 0.7%; Score 8; DB 4; Length 95;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88  
  
RESULT 31  
O97514 PRELIMINARY; PRT; 95 AA.  
AC O97514;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Pan troglodytes (Chimpanzee).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
OX NCBI\_TaxID=9598;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF010294; AAC67222.1; -.  
GO; GO:0016021; C:integral to membrane; IEA.  
GO; GO:0004872; F:receptor activity; IEA.  
GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.  
GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCR\_Rhodopsn.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER 95 95  
SQ SEQUENCE 95 AA; 9119 MW; 02BD95057095863C CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 95;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88  
  
RESULT 32  
Q9N249 PRELIMINARY; PRT; 95 AA.  
AC Q9N249;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Hylobates klossii (Kloss's gibbon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.  
OX NCBI\_TaxID=9598;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF010294; AAC67222.1; -.  
GO; GO:0016021; C:integral to membrane; IEA.  
GO; GO:0004872; F:receptor activity; IEA.  
GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.  
GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCR\_Rhodopsn.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER 95 95  
SQ SEQUENCE 95 AA; 9119 MW; 02BD95057095863C CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 95;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88  
  
RESULT 33  
Q9N248 PRELIMINARY; PRT; 95 AA.  
AC Q9N248;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Hylobates lar (Common gibbon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.  
OX NCBI\_TaxID=9580;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125663; AAF64378.1; -.  
GO; GO:0016021; C:integral to membrane; IEA.  
GO; GO:0004872; F:receptor activity; IEA.  
GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.  
GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCR\_Rhodopsn.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER 95 95  
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 95;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88  
  
RESULT 34

OX NCBI\_TaxID=9587;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125662; AAF64377.1; -.  
GO; GO:0016021; C:integral to membrane; IEA.  
GO; GO:0004872; F:receptor activity; IEA.  
GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.  
GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCR\_Rhodopsn.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER 95 95  
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 95;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88  
  
RESULT 33  
Q9N248 PRELIMINARY; PRT; 95 AA.  
AC Q9N248;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Hylobates lar (Common gibbon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.  
OX NCBI\_TaxID=9580;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125663; AAF64378.1; -.  
GO; GO:0016021; C:integral to membrane; IEA.  
GO; GO:0004872; F:receptor activity; IEA.  
GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.  
GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCR\_Rhodopsn.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER 95 95  
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 95;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88  
  
RESULT 34



```

O97516
ID O97516 PRELIMINARY; PRT; 95 AA.
AC O97516;
DT 01-MAY-1999 (TREMELrel. 10, Created)
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deindard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates."
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010296; AAC67224.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . ; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
FT SEQUENCE 95 AA; 9121 MW; ECF6C50B3A809D33 CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 6; Length 95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

O97520
ID O97520 PRELIMINARY; PRT; 95 AA.
AC O97520;
DT 01-MAY-1999 (TREMELrel. 10, Created)
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deindard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates."
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010300; AAC67228.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . ; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
FT SEQUENCE 95 AA; 9295 MW; 1C5CD4013CB365E CRC64;

```

```

O97518
ID O97518 PRELIMINARY; PRT; 95 AA.
AC O97518;
DT 01-MAY-1999 (TREMELrel. 10, Created)
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deindard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates."
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010298; AAC67226.1; -.

Query Match
Best Local Similarity 0.7%; Score 8; DB 6; Length 95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

O97518
ID O97518 PRELIMINARY; PRT; 99 AA.
AC O97518;
DT 01-MAY-1999 (TREMELrel. 10, Created)
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deindard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates."
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010298; AAC67226.1; -.

Query Match
Best Local Similarity 0.7%; Score 8; DB 6; Length 95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

```

```
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0004872; F: receptor activity; IEA.
DR GO: GO:0001584; F: rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P: G-protein coupled receptor protein signaling; IEA.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsin.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
FT NON TER 99
SQ SEQUENCE 99 AA; 9535 MW; D4CED9750184BC47 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 99;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLLV 1161
Db 85 LLLALLLV 92

RESULT 38
Q8E9P1 PRELIMINARY; PRT; 104 AA.
ID Q8E9P1;
AC Q8E9P1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cell division protein FtsL.
DE FtsL OR SO4226.
GN Shewanella oneidensis.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123 (2002).
DR EMBL; AS015855; AAN57198.1; -.
DR TIGR; SO4226; -.
DR InterPro: IPR007082; FtsL.
DR Pfam; PF04999; FtsL; 1.
KW Complete proteome.
SQ SEQUENCE 104 AA; 12187 MW; 408B210954032920 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 104;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLLV 1161
Db 23 LLLALLLV 30

RESULT 39
Q8SDS4 PRELIMINARY; PRT; 106 AA.
ID Q8SDS4;
AC Q8SDS4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Non-specific lipid transfer protein.
```

```
GN LTPI.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamids; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Hambyul; TISSUE=Leaf;
RA Hwang B.K., Jung H.W.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF208832; AAF23458.1; -.
DR HSSP; P23096; 1RZL.
DR GO; GO:0008289; F: lipid binding; IEA.
DR GO; GO:0006869; P: lipid transport; IEA.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR000528; Plant LTP.
DR Pfam; PF00234; tryptophan alpha amy1; 1.
DR PRINTS; PR00382; LIPIDTRANSF.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 106 AA; 11291 MW; 559C43238B9381C9 CRC64;

Query Match 0.7%; Score 8; DB 10; Length 106;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 197 GOIQGVV 204
Db 20 GOIQGVV 27

RESULT 40
Q8R1E9 PRELIMINARY; PRT; 112 AA.
ID Q8R1E9;
AC Q8R1E9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN C2A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Salivary gland;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024677; AAN24677.1; -.
DR MGD; MGI:2655426; C2a.
DR InterPro: IPR006038; Uteroglobin_supf.
KW Hypothetical protein.
SQ SEQUENCE 112 AA; 12371 MW; 69CE957825183F72 CRC64;

Query Match 0.7%; Score 8; DB 11; Length 112;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 LLLALLLV 1160
Db 5 LLLALLLV 12

RESULT 41
Q8JZX1 PRELIMINARY; PRT; 112 AA.
ID Q8JZX1;
AC Q8JZX1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Salivary androgen-binding protein gamma
DE subunit).
```

C2D OR ABPG.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/HeJ;  
 RA Strausberg R.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/HeJ;  
 RX MEDLINE=22680631; PubMed=12795612;  
 RA Karn R.C., Laukaitis C.M.;  
 RT "Characterization of two forms of mouse salivary androgen-binding  
 protein (ABP): implications for evolutionary relationships and ligand-  
 binding function.";  
 RL Biochemistry 42:7162-7170(2003).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/HeJ;  
 RA Laukaitis C.M., Dlouhy S.R., Karn R.C.;  
 RT "The mouse salivary androgen-binding protein (ABP) gene cluster on  
 chromosome 7: Characterization and evolutionary relationships.";  
 RL Mamm. Genome 0:0-0(2003).  
 DR EMBL; BC035222; AAP35222.1; -.  
 DR EMBL; AY325898; AAP94990.1; -.  
 DR MGD; MGI:2655745; C2d.  
 KW Hypothetical protein.  
 SQ SEQUENCE 112 AA; 12794 MW; F832B5868F4B3D53D CRC64;  
 Query Match 0.7%; Score 8; DB 11; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1153 LLLLLLV 1160  
 DB 5 LLLLLLV 12

RESULT 42  
 Q7TNY5 PRELIMINARY; PRT; 112 AA.  
 ID Q7TNY5  
 AC Q7TNY5  
 DT 01-OCT-2003 (TREMELrel. 25, Created)  
 DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Salivary androgen-binding protein beta subunit.  
 GN ABPG.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/HeJ;  
 RX MEDLINE=22680631; PubMed=12795612;  
 RA Karn R.C., Laukaitis C.M.;  
 RT "Characterization of two forms of mouse salivary androgen-binding  
 protein (ABP): implications for evolutionary relationships and ligand-  
 binding function.";  
 RL Biochemistry 42:7162-7170(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/HeJ;  
 RA Laukaitis C.M., Dlouhy S.R., Karn R.C.;  
 RT "The mouse salivary androgen-binding protein (ABP) gene cluster on  
 chromosome 7: Characterization and evolutionary relationships.";  
 RL Mamm. Genome 0:0-0(2003).  
 DR EMBL; AY325897; AAP94989.1; -.  
 SQ SEQUENCE 112 AA; 12383 MW; EB5F2C68DBF8D17C CRC64;

Query Match 0.7%; Score 8; DB 11; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1153 LLLLLLV 1160  
 DB 5 LLLLLLV 12

RESULT 43  
 Q7TNN0 PRELIMINARY; PRT; 112 AA.  
 ID Q7TNN0  
 AC Q7TNN0  
 DT 01-OCT-2003 (TREMELrel. 25, Created)  
 DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Salivary androgen-binding protein beta subunit.  
 GN ABPG.  
 OS Mus musculus (eastern European house mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=39442;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DBA/2J;  
 RX MEDLINE=22680631; PubMed=12795612;  
 RA Karn R.C., Laukaitis C.M.;  
 RT "Characterization of two forms of mouse salivary androgen-binding  
 protein (ABP): implications for evolutionary relationships and ligand-  
 binding function.";  
 RL Biochemistry 42:7162-7170(2003).  
 DR EMBL; AY293278; AAP44465.1; -.  
 SQ SEQUENCE 112 AA; 12411 MW; 818431A8CF8C029 CRC64;  
 Query Match 0.7%; Score 8; DB 11; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1153 LLLLLLV 1160  
 DB 5 LLLLLLV 12

RESULT 44  
 Q7TNW7 PRELIMINARY; PRT; 112 AA.  
 ID Q7TNW7  
 AC Q7TNW7  
 DT 01-OCT-2003 (TREMELrel. 25, Created)  
 DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Salivary androgen-binding protein gamma subunit.  
 GN ABPG.  
 OS Mus musculus (eastern European house mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=39442;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DBA/2J;  
 RX MEDLINE=22680631; PubMed=12795612;  
 RA Karn R.C., Laukaitis C.M.;  
 RT "Characterization of two forms of mouse salivary androgen-binding  
 protein (ABP): implications for evolutionary relationships and ligand-  
 binding function.";  
 RL Biochemistry 42:7162-7170(2003).  
 DR EMBL; AY293281; AAP44468.1; -.  
 SQ SEQUENCE 112 AA; 12546 MW; 9F955317D9116883 CRC64;  
 Query Match 0.7%; Score 8; DB 11; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1153 LLLLLLV 1160

```
Db      TIGR; DR2210; -.
KW      Hypothetical protein; Complete proteome
SQ      SEQUENCE 127 AA; 14314 MW; 2B3E79D3522A081 CRC64;

Query Match      0.7%; Score 8; DB 16; Length 127;
Best Local Similarity 100.0%; Pred.No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1152 GLLLLALL 1159
Db      7 GLLLLALL 14

RESULT 47
Q82U09 PRELIMINARY; PRT; 127 AA.
AC Q82U09
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 25, Last annotation update)
DE CRCB protein.
GN NEI704.
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19718 / IF0 14298;
RX MEDLINE=22586410; PubMed=12700255;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA Arciero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;
RT "Complete genome sequence of the ammonia-oxidizing bacterium and
RT obligate chemolithoautotroph Nitrosomonas europaea.";
RL J. Bacteriol. 185:2759-2773(2003).
DR EMBL; BX321862; CAD85615.1; -.
DR GO; GO:0016020; C-membrans; IEA.
DR InterPro; IPR003691; Camphor_CrbB.
DR Pfam; PF02537; CRCB; 1.
DR TIGRPFAMS; TIGR00494; crcb; 1.
KW Complete proteome.
SQ SEQUENCE 127 AA; 13419 MW; 55F30C8DD8AD7952 CRC64;

Query Match      0.7%; Score 8; DB 16; Length 127;
Best Local Similarity 100.0%; Pred.No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1147 GSTLGGLL 1154
Db      11 GSTLGGLL 18

RESULT 48
Q8C442 PRELIMINARY; PRT; 137 AA.
AC Q8C442
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Stimulated by retinoic acid gene 6.
GN STRA6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Hippocampus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
```

```

RT 60,770 full-length cDNAs.;
RL Nature 420:563-573(2002).;
DR EMBL; AK083120; BAC38769.1; -.
DR PIR; PT0714; PT0714.
DR MGI; MGI:107742; Strad.
SQ SEQUENCE 137 AA; 14573 MW; 27A54C974EF867CC CRC64;

Query Match      0.7%; Score 8; DB 11; Length 137;
Best Local Similarity 100.0%; Pred.No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLIV 1160
Db 64 LLLLLLIV 71

RESULT 49
Q61692 PRELIMINARY; PRT; 141 AA.
AC Q61692;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative heat stable antigen.
GN HSA-C.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBA X C57BL/6; TISSUE=Spleen;
RX MEDLINE=91209380; Pubmed=2019286;
RA Wenger R.H., Avane M., Bose R., Koehler G., Nielsen P.J.;
RT "The genes for a mouse hematopoietic differentiation marker called the
RT heat-stable antigen.";
RL Eur. J. Immunol. 21:1039-1046(1991).
DR EMBL; X55486; CAA39843.1; -.
DR PIR; S15785; S15785.
SQ SEQUENCE 141 AA; 15515 MW; E4BFB428ADC03C69 CRC64;

Query Match      0.7%; Score 8; DB 11; Length 141;
Best Local Similarity 100.0%; Pred.No. 30;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 12 GLLLLALL 19

RESULT 50
Q7TIM3 PRELIMINARY; PRT; 144 AA.
AC Q7TIM3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bradykinin-potentiating/C-type natriuretic peptide isoform 2
DE (Fragment).
OS Bothrops jararacussu (Jararacussu).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
CX NCBI_TaxID=8726;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Hayashi M.A.F., Lameu C., Radis-Baptista G., Yamane T.,
RA Camargo A.C.M.;
RT "Cloning and sequence analysis of a Bothrops jararacussu BPPs
RT precursor.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY310915; AAP83422.1; -.

```

---

```

FT NON TER 144 144
SQ SEQUENCE 144 AA; 15982 MW; B23FED5230A84910 CRC64;

Query Match      0.7%; Score 8; DB 13; Length 144;
Best Local Similarity 100.0%; Pred.No. 30;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 10 GLLLLALL 17

RESULT 51
Q8QG90 PRELIMINARY; PRT; 146 AA.
AC Q8QG90;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Bradykinin-potentiating protein.
OS Bothrops insularis (Island jararaca) (Queimada jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
CX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Junqueira-de-Azevedo I.L.M., Ho P.L.;
RT "A survey of gene expression and diversity in the venom glands of the
RT pitviper Bothrops insularis through the generation of Expressed
RT Sequence Tags (ESTs).";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF490532; AAM09691.1; -.
SQ SEQUENCE 146 AA; 16088 MW; F58244C4F24B1799 CRC64;

Query Match      0.7%; Score 8; DB 13; Length 146;
Best Local Similarity 100.0%; Pred.No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 10 GLLLLALL 17

RESULT 52
Q7XK40 PRELIMINARY; PRT; 161 AA.
AC Q7XK40;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE OSUNBA0044K18.5 protein.
GN OSUNBA0044K18.5.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
CX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,
RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,
RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Bing C.W.,
RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu P.H., Han B., Feng Q.,
RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Wu J., Yu Z.,
RA Chen L., Fan D.L., Wang Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,
RA Lu T.T., Zhang Y.J., Li Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,
RA Qian Y.M., Yifeng K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,
RA Zhang R.Q., Guan J.P., Hong G.F.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL606595; CAE05863.1; -.
SQ SEQUENCE 161 AA; 16300 MW; 0E8B2DF9C58ADC61 CRC64;

```

```
Query Match      0.7%; Score 8; DB 10; Length 161;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1156 LALLVLLAL 1163
Db 136 LALLVLLAL 143

RESULT 53
Q98MM8 PRELIMINARY; PRT; 174 AA.
ID Q98MM8
AC Q98MM8
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
GN Hypothetical protein mlr0512.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Katanabe A., Ideawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP002995; BAB48085.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 174 AA; 20270 MW; E91E12118C11022F CRC64;

Query Match      0.7%; Score 8; DB 16; Length 174;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1090 LRLSLKALK 1097
Db 38 LRLSLKALK 45

RESULT 54
P79786 PRELIMINARY; PRT; 181 AA.
ID P79786
AC P79786
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Thrombospondin-1 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97209337; PubMed=9056637;
RA Tucker R.F., Hagios C., Chiquet-Ehrismann R., Lawler J.;
RT "In situ localization of thrombospondin-1 and thrombospondin-3
RT transcripts in the avian embryo.";
RL Dev. Dyn. 208:326-337(1997).
DR EMBL; U76994; AAB19208.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003367; tsep_3.
DR Pfam; PF02412; tsep_3; 8.

Query Match      0.7%; Score 8; DB 13; Length 181;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159
Db 10 GLLLLALL 17

RESULT 56
Q90Y11 PRELIMINARY; PRT; 181 AA.
ID Q90Y11
AC Q90Y11
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Bradykinin potentiating peptide and C-type natriuretic peptide isoform
DE 2.
OS Crotalus durissus terrificus (South American rattlesnake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Crotalus.
OX NCBI_TaxID=8732;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Hayashi M.A.F., Radis-Baptista G., Barbosa S.R., Yamane T.,
RA Camargo A.C.M.;
```

RT \*Crotalus durissus terrificus bradykinin-potentiating peptide and C-  
RT type natriuretic peptide precursor isoform2.\*;  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF308594; AAL09427.1; -;  
DR GO: GO:0005576; C:extracellular; IEA.  
DR GO: GO:0005179; F:hormone activity; IEA.  
DR InterPro: IPR000663; Natl\_peptide.  
DR Pfam: PF00212; ANP; 1.  
DR PRINTS: PR00710; NATPEPTIDES.  
DR SMART: SM00183; NAT\_PEP; 1.  
DR PROSITE: PS00263; NATRIURETIC\_PEP\_TIDE; 1.  
SQ SEQUENCE 181 AA; 18507 MW; 9B2B95D38AA5FF27 CRC64;

Query Match 0.7%; Score 8; DB 13; Length 181;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159  
DB 10 GLLLLALL 17

RESULT 57

Q8U4N4 PRELIMINARY; PRT; 222 AA.  
ID Q8U4N4  
AC Q8U4N4  
DT 01-JUN-2002 (TREMELrel. 21, Created)  
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
DE Hypothetical protein PF0047.  
GN PF0047.  
OS Pyrococcus furiosus.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Pyrococcus.  
OX NCBI\_TaxID=2261;  
RN [1]  
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;  
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;  
RT "The complete sequence of the Pyrococcus furiosus genome.";  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB010131; AAL80171.1; -;  
RW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 222 AA; 24556 MW; 721D79ED547A0F13 CRC64;

Query Match 0.7%; Score 8; DB 17; Length 222;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 512 RGRVYVE 519  
DB 106 RGRVYVE 113

RESULT 58

Q7V5L4 PRELIMINARY; PRT; 224 AA.  
ID Q7V5L4  
AC Q7V5L4  
DT 01-OCT-2003 (TREMELrel. 25, Created)  
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE Conserved hypothetical membrane protein.  
GN P3H9.20 OR P3H1539.  
OS Prochlorococcus marinus (strain MIT 9313).  
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
OC Prochlorococcus.  
OX NCBI\_TaxID=74547;  
RN [1]  
RC MEDLINE=22825698; PubMed=12917642;  
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,  
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,  
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,

RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,  
RA Webb E.A., Zinser E.R., Chisholm S.W.;  
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic  
RT niche differentiation.";  
RL Nature 424:1042-1047(2003).  
DR EMBL: BX572099; CA321714.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 224 AA; 24172 MW; E1CD9F149C279A29 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 224;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1158 LVLALRK 1165  
DB 190 LVLALRK 197

RESULT 59

OS7835 PRELIMINARY; PRT; 224 AA.  
ID OS7835  
AC OS7835  
DT 01-AUG-1998 (TREMELrel. 07, Created)  
DT 01-AUG-1998 (TREMELrel. 07, Last sequence update)  
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
DE Hypothetical protein PH0087.  
GN PH0087  
OS Pyrococcus horikoshii.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Pyrococcus.  
OX NCBI\_TaxID=53953;  
RN [1]  
RC STRAIN=OT3;  
RX MEDLINE=98344137; PubMed=9679194;  
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,  
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
RA Masuchi Y., Shizuya H., Kikuchi H.;  
RT "Complete sequence and gene organization of the genome of a hyper-  
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA Res. 5:55-76(1998).  
DR EMBL: AP000001; BAA29156.1; -;  
DR PIR: E71228; E71228.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 224 AA; 25134 MW; 9CC56212064558E0 CRC64;

Query Match 0.7%; Score 8; DB 17; Length 224;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 512 RGRVYVE 519  
DB 105 RGRVYVE 112

RESULT 60

P74604 PRELIMINARY; PRT; 226 AA.  
ID P74604  
AC P74604  
DT 01-FEB-1997 (TREMELrel. 02, Created)  
DT 01-FEB-1997 (TREMELrel. 02, Last sequence update)  
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
DE Hypothetical protein slr1571.  
GN SLR1571.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RC MEDLINE=97061201; PubMed=8905231;  
RX SEQUENCE FROM N.A.



```

RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.,
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.",
RL DNA Res. 3:109-136(1996).
DR EMBL; D90916; BAA18712.1; -.
DR FIRM; S76800;
DR InterPro; IPR002838; DUF124.
DR Pfam; PF01587; DUF124; 1.
DR ProDom; PD013634; DUF124; 1.
DR TIGRfam; TIGR00266; TIGR00266; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 226 AA; 25003 MW; 29C5C137C128DAC0 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 226;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1073 RLVPNQEI 1080
DB 183 RLVPNQEI 190

RESULT 61
Q882P2 PRELIMINARY; PRT; 227 AA.
AC Q882P2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Conserved hypothetical protein.
GN PSPT02584.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidson T.,
RA White O., Fraser C., Collier A.;
RT "Complete sequence of Pseudomonas syringae.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; A2016865; AAC056088.1; -.
DR TIGR; PSPT02584; -.
DR InterPro; IPR002489; DUF14.
DR Pfam; PF01493; GXKG; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 227 AA; 24036 MW; AB3458F347BEEF145 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 227;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 547 ASVRDLNQ 554
DB 9 ASVRDLNQ 16

RESULT 62
Q9AJX4 PRELIMINARY; PRT; 231 AA.
ID Q9AJX4;
AC Q9AJX4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```

```

DE Putative oxidoreductase.
OS SC01803 OR SC133.02.
GN Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Harris D.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kisser H.M., Denapaire D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=21996410; PubMed=12000953;
RA Thomson N.R., James K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kisser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC - SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
DR HSSP; AL939110; CAC28543.1; -.
DR HSSP; 064105; 10AA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 231 AA; 23789 MW; 6AEDB40306B94085 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 231;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 790 DLVLIDARS 797
DB 27 DLVLIDARS 34

RESULT 63
Q98GF8 PRELIMINARY; PRT; 235 AA.
ID Q98GF8;
AC Q98GF8;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ATP-binding component of phosphate transport, HisP-like
DE nucleotide binding protein, PhnL.
GN MLR3349.

```

```

OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Matanabe A., Igesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RL Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AF003001; BAB50258.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_TRANSPORTER.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD00006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 235 AA; 25098 MW; 6FA8F03F3F827361 CRC64;

Query Match          0.7%; Score 8; DB 16; Length 235;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 EDGLIDLA 623
DB 73 EDGLIDLA 80

RESULT 64
Q20220 PRELIMINARY; PRT; 245 AA.
AC Q20220;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein F40F4.7.
GN F40F4.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RT "The sequence of C. elegans cosmid F40F4.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL; U04020; AAK84526.3; -.
DR WormPep; F40F4.7; CE28316.
DR GO; GO:0008080; F:N-acetyltransferase activity; IEA.
DR InterPro; IPR00182; GCM5acetyl_trans.
DR Pfam; PF00583; Acetyltrans; 1.
KW Hypothetical protein.
SQ SEQUENCE 245 AA; 27893 MW; AS3289FBB2F91E75 CRC64;

Query Match          0.7%; Score 8; DB 5; Length 245;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 557 YNDVVVGA 564
DB 141 YNDVVVGA 148

RESULT 65
Q8Y1R9 PRELIMINARY; PRT; 257 AA.
AC Q8Y1R9;
DT 01-VAR-2002 (TrEMBLrel. 20, Created)
DT 01-VAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein RSC0620.
GN RSC0620 OR RS01521.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Signier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646060; CAD14150.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000572; Oxidored_molyp.
DR Pfam; PF00174; oxidored_molyp; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 257 AA; 29260 MW; DAE569EE1572113D CRC64;

Query Match          0.7%; Score 8; DB 16; Length 257;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1149 TLGGILLL 1156
DB 37 TLGGILLL 44

RESULT 66
Q8QG91 PRELIMINARY; PRT; 265 AA.
AC Q8QG91;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Bradykinin-potentiating/c-type natriuretic protein.
OS Bothrops insularis (Island jararaca) (Queimada jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE=Venom gland;
RA Junqueira-de-Azevedo I.L.M., Ho P.L.;
RT "A survey of gene expression and diversity in the venom glands of the
RT pitviper Bothrops insularis through the generation of Expressed
RT Sequence Tags (ESTs)."; to the EMBL/GenBank/DBJ databases.
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF490531; AAM09690.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR000663; Natr_peptide.
DR Pfam; PF00212; ANP; 1.
DR PRINTS; PR00710; NATPEPTIDES.
DR SMART; SM00183; NAT_PEP; 1.
DR PROSITE; PS00263; NATRIURETIC_PEPTIDE; 1.
SQ SEQUENCE 265 AA; 27763 MW; 0EAE1408B42358BE CRC64;

Query Match 0.7%; Score 8; DB 13; Length 265;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 10 GLLLLALL 17

RESULT 67
Q9PW56 PRELIMINARY; PRT; 265 AA.
AC Q9PW56;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Bradykinin-potentiating/C-type natriuretic peptide.
OS Bothrops jararaca (Jaratoca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8724;
RN [1] SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Hayashi M.A.F., Murbach A.F., Camargo A.C.M.;
RT "The precursor of C-type natriuretic peptide of snake brain contains
RT angiotensin converting enzyme inhibitors, specific for the C-catalytic
RT site.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF171670; AAD51326.2; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR000663; Natr_peptide.
DR Pfam; PF00212; ANP; 1.
DR PRINTS; PR00710; NATPEPTIDES.
DR SMART; SM00183; NAT_PEP; 1.
DR PROSITE; PS00263; NATRIURETIC_PEPTIDE; 1.
SQ SEQUENCE 265 AA; 27763 MW; 8E9AEC976CCD439 CRC64;

Query Match 0.7%; Score 8; DB 13; Length 265;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 10 GLLLLALL 17

RESULT 68
Q861M5 PRELIMINARY; PRT; 272 AA.
AC Q861M5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MHC class I antigen (Fragment).

```

```

OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1] SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=107;
RA Chung C., Leib S.R., Fraser D.G., Ellis S.A., McGuire T.C.;
RT "Sequence-based typing of horse classical MHC class I genes.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY176095; AAO18112.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006955; F:immune response; IEA.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG_c1.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR SMART; SM00407; IGC1; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
FT NON_TER 1
FT NON_TER 272
SQ SEQUENCE 272 AA; 30576 MW; 8EC9E0C183D319FD CRC64;

Query Match 0.7%; Score 8; DB 7; Length 272;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 VLLGAVGA 387
DB 232 VLLGAVGA 239

RESULT 69
Q861K3 PRELIMINARY; PRT; 272 AA.
AC Q861K3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MHC class I antigen (Fragment).
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1] SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=130;
RA Chung C., Leib S.R., Fraser D.G., Ellis S.A., McGuire T.C.;
RT "Sequence-based typing of horse classical MHC class I genes.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY176118; AAO18134.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006955; F:immune response; IEA.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG_c1.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR PRODOM; PD000050; MHC_I; 1.
DR SMART; SM00407; IGC1; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
FT NON_TER 1
FT NON_TER 272
SQ SEQUENCE 272 AA; 30220 MW; F7FC66FFB06558CF CRC64;

Query Match 0.7%; Score 8; DB 7; Length 272;

```

```

Best local Similarity 100.0%; Pred. No. 53;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 VLLGAVGA 387
Db 232 VLLGAVGA 239

RESULT 70
Q8P8NG PRELIMINARY; PRT; 278 AA.
AC Q8P8NG
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE ABC transporter sugar permease.
GN LAGC OR KCC204.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannaval F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Noon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.P., Takita M.A., Tamura R.E., Teixeira B.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White P.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AE012328; AM41484.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD_transp.
DR Pfam; PF00528; BPD_transp; 1.
RW Complete proteome.
SQ SEQUENCE 278 AA; 30673 MW; FB45244D924B7361 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 278;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLAL 1158
Db 17 GGLLLAL 24

RESULT 71
Q8TPV4 PRELIMINARY; PRT; 290 AA.
AC Q8TPV4
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Uncharacterized domain specific for M.kandleri, MK-34 family, a.
GN MK1344.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.

```

```

OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozlovskiy S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL; AF010427; AA02557.1; -.
RW Complete proteome.
SQ SEQUENCE 290 AA; 32754 MW; 706E6883976ADF91 CRC64;

Query Match 0.7%; Score 8; DB 17; Length 290;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 268 VITDGESH 275
Db 161 VITDGESH 168

RESULT 72
Q94K03 PRELIMINARY; PRT; 299 AA.
AC Q94K03
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Unknown protein.
GN A_IQ02N01.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF370497; AA43874.1; -.
RW Complete proteome.
SQ SEQUENCE 299 AA; 34028 MW; BB2AB03C4A184D3E CRC64;

Query Match 0.7%; Score 8; DB 10; Length 299;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
Db 44 LLLALLLV 51

RESULT 73
Q98BD8 PRELIMINARY; PRT; 307 AA.
AC Q98BD8
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein ml15620.
GN ML15620.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;

```

```

RN  SEQUENCE FROM N.A.
RP  STRAIN=NAFF303099;
RC  MEDLINE=21082930; PubMed=11214968;
RA  Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RX  Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA  Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA  Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA  Takeuchi C., Yamada M., Tabata S.;
RT  "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT  Mesorhizobium loti.";
RL  DNA Res. 7:331-338(2000).
DR  EMBL; AP003007; BAB52034.1; -.
DR  InterPro: IPR002781; DUF81.
DR  Pfam; PF01925; DUF81; 1.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 307 AA; 31951 MW; 938FFC1D38C4AC16 CRC64;

Query Match      0.7%; Score 8; DB 16; Length 307;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1155 LLALLVLA 1162
DB  275 LLALLVLA 282
|||||

RESULT 74
Q92885 PRELIMINARY; PRT; 308 AA.
AC  Q92885;
DT  01-DEC-2001 (TrEMBLrel. 19, Created)
DT  01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  Hypothetical transmembrane protein SMC02232.
GN  R00337 OR SMC02232.
OS  Rhizobium meliloti (Sinorhizobium meliloti).
OC  Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC  Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX  NCBI_TaxId=382;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=1021;
RX  MEDLINE=21396507; PubMed=11481430;
RA  Capela D., Barloy-Hubler F., Guzy J., Bothe G., Ampe F., Batut J.,
RA  Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA  Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA  Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA  Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT  "Analysis of the chromosome sequence of the legume symbiont
RT  Sinorhizobium meliloti strain 1021.";
RL  Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR  EMBL; AL591784; CAC45109.1; -.
DR  InterPro: IPR002781; DUF81.
DR  Pfam; PF01925; DUF81; 1.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 308 AA; 32465 MW; C86497E6E161BC0D CRC64;

Query Match      0.7%; Score 8; DB 16; Length 308;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1155 LLALLVLA 1162
DB  275 LLALLVLA 282
|||||

RESULT 75
Q8CB84 PRELIMINARY; PRT; 323 AA.
ID  Q8CB84
AC  Q8CB84;
DT  01-MAR-2003 (TrEMBLrel. 23, Created)
DT  01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

```

```

DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  Integrin alpha 2 (Fragment).
GN  ITGA2.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxId=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Bone;
RX  MEDLINE=22354683; PubMed=12466851;
RA  The FANTOM Consortium,
RA  "Analysis of the mouse transcriptome based on functional annotation of
RT  60,770 full-length cDNAs.";
RL  Nature 420:563-573(2002).
DR  EMBL; AK036572; BAC29485.1; -.
DR  MGD; MGI:96600; Itga2.
DR  GO; GO:0008305; C:integrin complex; IEA.
DR  GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR  GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR  InterPro: IPR000413; Integrin_alpha.
DR  InterPro: IPR002035; VWF_A.
DR  Pfam; PF00092; vwa; 1.
DR  PRINTS; PR00453; VWFADOMAIN.
DR  SMART; SM00191; Int_alpha; 1.
DR  SMART; SM00327; VWA; 1.
DR  PROSITE; PS50234; VWFA; 1.
FT  NON TER 323 323
SQ  SEQUENCE 323 AA; 35025 MW; 05A456A44283724A CRC64;

Query Match      0.7%; Score 8; DB 11; Length 323;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  71 GDVYKCPV 78
DB  75 GDVYKCPV 82
|||||

Search completed: June 24, 2004, 18:06:14
Job time : 63 secs

```



OM protein - protein search, using sw model

Run on: June 24, 2004, 17:56:51 ; Search time 65 Seconds  
(without alignments)  
5164.099 Million cell updates/sec

Title: US-09-980-403-2  
Perfect score: 1188  
Sequence: 1 MDLPRGLVAVALLSLWPGLT.....FRSARRRPPGLDTPKVLK 1188

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : A\_Geneseq\_29Jan04.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001s.\*
- 5: Geneseq2002s.\*
- 6: Geneseq2003as.\*
- 7: Geneseq2003Bs.\*
- 8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description         |
|------------|-------|---------------|--------|-------|---------------------|
| 1          | 1188  | 100.0         | 1188   | 4     | AAB30929 Amino aci  |
| 2          | 928   | 78.1          | 1189   | 4     | ABG12949 Novel hum  |
| 3          | 926   | 77.9          | 1188   | 4     | Aau14231 Human nov  |
| 4          | 926   | 77.9          | 1188   | 4     | AAB50085 Human nov  |
| 5          | 926   | 77.9          | 1188   | 5     | Aau10551 Human A25  |
| 6          | 911   | 76.7          | 1188   | 4     | Aau14467 Human nov  |
| 7          | 911   | 76.7          | 1188   | 7     | ADP09956 Novel pro  |
| 8          | 911   | 76.7          | 1189   | 3     | AAB25582 ITG11 pr   |
| 9          | 911   | 76.7          | 1189   | 6     | ABR58364 Human nov  |
| 10         | 911   | 76.7          | 1189   | 6     | ADA27054 Human nov  |
| 11         | 911   | 76.7          | 1189   | 7     | ADP63570 Human pro  |
| 12         | 911   | 76.7          | 1189   | 8     | ADP65854 Human nov  |
| 13         | 810   | 68.2          | 1034   | 3     | AAB25590 Protein e  |
| 14         | 810   | 68.2          | 1034   | 6     | ADA27062 Human nov  |
| 15         | 810   | 68.2          | 1034   | 8     | ADP6592 Human hum   |
| 16         | 733   | 63.4          | 1120   | 6     | ABR58365 Human NOV  |
| 17         | 489   | 41.2          | 707    | 4     | Aau19663 Human nov  |
| 18         | 489   | 41.2          | 707    | 5     | ABP47893 Human pol  |
| 19         | 489   | 41.2          | 707    | 7     | ADP10845 Human ext  |
| 20         | 193   | 16.2          | 193    | 5     | Aau176854 Human int |
| 21         | 190   | 16.0          | 193    | 5     | Aau176863 Human int |
| 22         | 122   | 10.3          | 1188   | 4     | AAB50087 Murine A2  |
| 23         | 122   | 10.3          | 1188   | 5     | Aau10552 Murine A2  |
| 24         | 104   | 8.8           | 360    | 7     | ADP08585 Novel pro  |
| 25         | 89    | 7.5           | 109    | 5     | ABG66673 Human nov  |



99 8 0.7 17 2 AAW5098 Aaw5098 Canine be  
100 8 0.7 17 2 AAW72832 Aaw72832 Canine al

ALIGNMENTS

RESULT 1

AAB30929 standard; protein; 1188 AA.

XX AAB30929;

02-APR-2001 (first entry)

DE Amino acid sequence of a human alpha1 integrin chain.

XX Human; integrin; alpha1 subunit; fibroblast; muscle cell; chondrocyte;  
KW osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;  
KW wound healing; trauma; rheumatoid arthritis; osteoarthritis;  
KW osteoporosis; cartilage damage; bone damage; cartilage.

OS Homo sapiens.

XX Key Location/Qualifiers

PH Peptide 1..22

FT /note= "signal peptide"

FT 951..972

FT /note= "leucine zipper"

FT 1142..1184

FT /note= "transmembrane domain"

XX W0200075187-A1.

XX 14-DEC-2000.

XX 31-MAY-2000; 2000WO-S2001135.

XX 03-JUN-1999; 99SE-00002056.

XX (ACTI-) ACTIVE BIOTECH AB.

XX Gullberg D;

XX WPI; 2001-071061/08.

XX N-PSDS; AAC86871.

XX Integrin subunit alpha 11 or integrin heterodimer comprising subunit  
PT alpha 11 in association with subunit beta, useful for treating muscle  
PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.

XX Disclosure; Fig 2a-c; 79pp; English.

XX The present sequence represents a human integrin subunit, designated  
CC alpha11. The alpha11 polynucleotide and polypeptide are useful as markers  
CC of cell target molecules, such as fibroblasts, muscle cells,  
CC chondrocytes, osteoblasts, mesenchymally derived cells or stem cells.  
CC They are also used for determining the differential stage of cells during  
CC differentiation, development in pathological conditions, in tissue  
CC regeneration, in transplantation or in therapeutic and physiological  
CC repair of tissues. The pathological conditions involving subunit alpha11  
CC are selected from damage of cells, muscle dystrophy, fibrosis, wound  
CC healing, trauma, rheumatoid arthritis, osteoarthritis and osteoporosis,  
CC damage of cartilage and bone and cartilage and bone diseases. The  
CC polypeptide is useful for detecting the formation of cartilage during  
CC embryonic development, for detecting physiological therapeutic repair of  
CC cartilage and muscle, for selection and analysis, or for sorting,  
CC isolating or purification of chondrocytes and muscle cells, for detecting  
CC regeneration of cartilage or chondrocytes during transplantation of  
CC cartilage or chondrocytes during transplantation of cartilage or  
CC chondrocytes, respectively, or of muscle or muscle cells during  
CC transplantation of muscle or muscle cells, respectively, and for studies  
CC of differentiation or chondrocytes or muscle cells

XX Sequence 1188 AA;  
SQ

Query Match 100.0%; Score 1188; DB 4; length 1188;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLPRGLVAVNALSILWPGFTDTFNDTRKPRVTPGSRATFAFGYTTQOHDISGNKWLAVCA 60  
DB 1 MDLPRGLVAVNALSILWPGFTDTFNDTRKPRVTPGSRATFAFGYTTQOHDISGNKWLAVCA 60  
QY 61 PLETNGYQKTDVYKCPVIHGNCTKLNLRVTLNVNRSKONRGLSLATNPKNSFLA 120  
DB 61 PLETNGYQKTDVYKCPVIHGNCTKLNLRVTLNVNRSKONRGLSLATNPKNSFLA 120  
QY 121 CSPLASHCEGSSYTTGMCNRVNSFRSKTVAPALQRCQTTMDIVLDGNSYTPWVE 180  
DB 121 CSPLASHCEGSSYTTGMCNRVNSFRSKTVAPALQRCQTTMDIVLDGNSYTPWVE 180  
QY 181 VQHFILNLIKXYIGPGQIQGVQYGVGVVHFEFLNDYRSVKDVEAASHIEQGGTET 240  
DB 181 VQHFILNLIKXYIGPGQIQGVQYGVGVVHFEFLNDYRSVKDVEAASHIEQGGTET 240  
QY 241 RTAFGIFARSEAFQGGKRGAKKVMIVITGESHSDSPLEKVIQOESRDNVTRAVAVL 300  
DB 241 RTAFGIFARSEAFQGGKRGAKKVMIVITGESHSDSPLEKVIQOESRDNVTRAVAVL 300  
QY 301 GYNRRGINPETFLNEIKYIASDDPDKHFFNTVDEAALKDIVDALGDRIFSLGNTKNET 360  
DB 301 GYNRRGINPETFLNEIKYIASDDPDKHFFNTVDEAALKDIVDALGDRIFSLGNTKNET 360  
QY 361 SFGLEMSQTGFSSHVVEDGVLGAVGAYDNGAVLKTSGKVIPLRESYLKEFPEELKN 420  
DB 361 SFGLEMSQTGFSSHVVEDGVLGAVGAYDNGAVLKTSGKVIPLRESYLKEFPEELKN 420  
QY 421 HGAYLGYTVTVSVSSROGRVYVAGAPRNHTCKVILFTMHNNRSITIHQAMGQQIGSYF 480  
DB 421 HGAYLGYTVTVSVSSROGRVYVAGAPRNHTCKVILFTMHNNRSITIHQAMGQQIGSYF 480  
QY 481 GSEITSDIDGQVTDVLLVGAPMYFNEGRGRGVYVYELRQNRVYNGTLKDSHYQNA 540  
DB 481 GSEITSDIDGQVTDVLLVGAPMYFNEGRGRGVYVYELRQNRVYNGTLKDSHYQNA 540  
QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFRGSILKTPKORITASELATG 600  
DB 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFRGSILKTPKORITASELATG 600  
QY 601 LQVFGCSIHQQLDLNEDGLIDLAVGALGNVILSRPVVQINASLHFPSPKINIFHRDCK 660  
DB 601 LQVFGCSIHQQLDLNEDGLIDLAVGALGNVILSRPVVQINASLHFPSPKINIFHRDCK 660  
QY 661 RSGRDATCLAAFLCFTPIFLAPHFTQTTVIGIRYNATMDERRYTTPRAHLDGGDRFTNRAV 720  
DB 661 RSGRDATCLAAFLCFTPIFLAPHFTQTTVIGIRYNATMDERRYTTPRAHLDGGDRFTNRAV 720  
QY 721 LLSGGQLCERINFVLDATDVVKPVTFSVEYSLEDDPHGPMDDGWPTTLRVSVFPWNG 780  
DB 721 LLSGGQLCERINFVLDATDVVKPVTFSVEYSLEDDPHGPMDDGWPTTLRVSVFPWNG 780  
QY 781 CNEDEHCVDPDLVDARSOLPTAMEYQVRLRPAQDCSAYTLSPDTTTPPIESTQRVAV 840  
DB 781 CNEDEHCVDPDLVDARSOLPTAMEYQVRLRPAQDCSAYTLSPDTTTPPIESTQRVAV 840  
QY 841 EATLNRGENAVSTVLNISQSANLQFASLIQKEDSGSIECVNERRRLOKQVNSYPPF 900  
DB 841 EATLNRGENAVSTVLNISQSANLQFASLIQKEDSGSIECVNERRRLOKQVNSYPPF 900  
QY 901 RAKAKVAFRLDSEFSKSIPLHLEITELAGDSNFRDSTKEDNVAPLPHLYEADVLPT 960  
DB 901 RAKAKVAFRLDSEFSKSIPLHLEITELAGDSNFRDSTKEDNVAPLPHLYEADVLPT 960  
QY 961 RSSSLSHVEVKNLSLSEVDGIGPPFSCIFRQNLGLPPIHGMKMTIPIATRSNRL 1020  
DB 961 RSSSLSHVEVKNLSLSEVDGIGPPFSCIFRQNLGLPPIHGMKMTIPIATRSNRL 1020

Db 961 RSSLSHVEVKLNSLERYDGIQPPFSCIFRQNLGLFPPIHGMKTIPIATRSNRL 1020  
 QY 1021 KLRDLTDEANTSCNINWNGSTERYPTPVEEDLRRAPQNLNHSNDVVSINCNIRLVPNCEI 1080  
 Db 1021 KLRDLTDEANTSCNINWNGSTERYPTPVEEDLRRAPQNLNHSNDVVSINCNIRLVPNCEI 1080  
 QY 1081 NFHLGNLWLSLKALKYKSKIMVNAALQRFPHSPPIFREEDPSRQIEFISKOEDMQV 1140  
 Db 1081 NFHLGNLWLSLKALKYKSKIMVNAALQRFPHSPPIFREEDPSRQIEFISKOEDMQV 1140  
 QY 1141 PIWIVGSTLGLLLALLLALRLKLGPFRRARRRREGLDPTPKVLE 1188  
 Db 1141 PIWIVGSTLGLLLALLLALRLKLGPFRRARRRREGLDPTPKVLE 1188

## RESULT 2

ABGI12949  
 ID ABGI12949 standard; protein; 1189 AA.  
 XX AC ABGI12949;  
 XX 13-FEB-2002 (first entry)  
 DT Novel human diagnostic protein #12940.  
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX Homo sapiens.  
 OS WO200175067-A2.  
 PN 11-OCT-2001.  
 PD 30-MAR-2001; 2001WO-US008631.  
 PF 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 PA Drmanac RT, Liu C, Tang YT;  
 PI WPI; 2001-639362/73.  
 DR N-PSDB; AAS77136.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 PS Claim 20; SEQ ID NO 43308; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1189 AA;  
 Query Match 78.13; Score 928; DB 4; Length 1189;  
 Best Local Similarity 99.94; Pred. No. 0;  
 Matches 1028; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MDLPRGLVAVWALSLSLWPGFTDTFNMDTRKPRVPGSRTAFYGYTVQOEDISGNKWLWVGA 60  
 Db 1 MDLPRGLVAVWALSLSLWPGFTDTFNMDTRKPRVPGSRTAFYGYTVQOEDISGNKWLWVGA 60  
 QY 61 PLETNGYQKTDGVYKCPVTHGNCTKLNIGRVTLNSVSKRKNRKLGLSLATNPKNSFLA 120  
 Db 61 PLETNGYQKTDGVYKCPVTHGNCTKLNIGRVTLNSVSKRKNRKLGLSLATNPKNSFLA 120  
 QY 121 CSPLWSHECGSSYYTTGKCSRVSNSNFRSKTVAPALQRCQTYMDIVIVLQGSNSIYPWVE 180  
 Db 121 CSPLWSHECGSSYYTTGKCSRVSNSNFRSKTVAPALQRCQTYMDIVIVLQGSNSIYPWVE 180  
 QY 181 VQHFLINILKFFYIGPQIQGVVQVQGEDVHFEHLNDYRSVKDVVAAASHIBORGTTET 240  
 Db 181 VQHFLINILKFFYIGPQIQGVVQVQGEDVHFEHLNDYRSVKDVVAAASHIBORGTTET 240  
 QY 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSHSDSPDLKVIQQSERDUNTRYAVAVL 300  
 Db 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSHSDSPDLKVIQQSERDUNTRYAVAVL 300  
 QY 301 GYNRRGINPETFLNEIKYIASDDPKHFFNVDEAALKDIVDALGDRIFSLGNTNKNET 360  
 Db 301 GYNRRGINPETFLNEIKYIASDDPKHFFNVDEAALKDIVDALGDRIFSLGNTNKNET 360  
 QY 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDWNGAVLAKETSAGKVIPLRESYIKPEPELKN 420  
 Db 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDWNGAVLAKETSAGKVIPLRESYIKPEPELKN 420  
 QY 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTGKVIPLFTMNNRSLTIHQAMRGQIGSYF 480  
 Db 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTGKVIPLFTMNNRSLTIHQAMRGQIGSYF 480  
 QY 481 GSEITSVDIDGCVTDVLLGAPMYNEGERGKVVYVELRONEFVYNGTLKQSHSYQNA 540  
 Db 481 GSEITSVDIDGCVTDVLLGAPMYNEGERGKVVYVELRONEFVYNGTLKQSHSYQNA 540  
 QY 541 RFGSSIASVRDLNQDSYNDVVVGAPLEDNHAGAIYIFHGPGRGSLTKPKQRTASELATG 600  
 Db 541 RFGSSIASVRDLNQDSYNDVVVGAPLEDNHAGAIYIFHGPGRGSLTKPKQRTASELATG 600  
 QY 601 LQYFGCSIHQQLDINEDGLIDLAVGALGNVILWSRPVQINASLHFEPSKINIHFHDCCK 660  
 Db 601 LQYFGCSIHQQLDINEDGLIDLAVGALGNVILWSRPVQINASLHFEPSKINIHFHDCCK 660  
 QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYTPRAHLDGGDRFTNRVAV 720  
 Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYTPRAHLDGGDRFTNRVAV 720  
 QY 721 LLSSGQELCERINPHVLDYADYKVPVTSVEYSLEDDPHGFMLDGGMPTTLRVSPFWNG 780  
 Db 721 LLSSGQELCERINPHVLDYADYKVPVTSVEYSLEDDPHGFMLDGGMPTTLRVSPFWNG 780  
 QY 781 CNDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSATYLSFDTTVPIIESTRQAVAV 840  
 Db 781 CNDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSATYLSFDTTVPIIESTRQAVAV 840  
 QY 841 EATLENRGENAYSTVLNISQANLQFASLQKEDSDGSIKESVNEERLQKQVCNVSYPFF 900  
 Db 841 EATLENRGENAYSTVLNISQANLQFASLQKEDSDGSIKESVNEERLQKQVCNVSYPFF 900  
 QY 901 RAKAKVAPRLDSPEKSIPLHLEIEILAAGSDNEDSTKEDNVAPLRFHLKYEADVLF 960  
 Db 901 RAKAKVAPRLDSPEKSIPLHLEIEILAAGSDNEDSTKEDNVAPLRFHLKYEADVLF 960  
 QY 961 RSSLSHVEVKLNSLERYDGIQPPFSCIFRQNLGLFPPIHGMKTIPIATRSNRL 1020

|||||  
 961 RSSLSHYEVLKSSLRDYGIGPPFSCIPRIQNLGLFPIHGMKTIPIATRSNRL 1020  
 1021 KLKDFLTDE 1029  
 1021 KLKDFLTDE 1029  
 RESULT 3  
 AAU14231  
 ID AAU14231 standard; protein; 1188 AA.  
 AC AAU14231;  
 DT 24-OCT-2001 (first entry)  
 XX Human novel protein #102.  
 DE  
 XX Human; novel protein; Antianemic; osteopathic; antiinflammatory;  
 KW immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic;  
 KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;  
 KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;  
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
 KW tissue regeneration; immune disorder.  
 XX  
 OS Homo sapiens.  
 PN W0200155437-A2.  
 XX 02-AUG-2001.  
 XX 25-JAN-2001; 2001WO-08002623.  
 XX 25-JAN-2000; 2000US-00491404.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C. Drmanac RT;  
 XX WPI; 2001-451939/48.  
 XX N-PSDB; AAS22536.  
 XX Isolated polypeptides useful for treating anti-inflammatory diseases,  
 XX nervous system disorders, and for regenerating bone and cartilage.  
 XX Example 4; Page 578-581; 894pp; English.  
 XX The invention relates to polynucleotides encoding novel human proteins or  
 XX their active domains. The polypeptides, polynucleotides and antibodies  
 XX raised against the polypeptides are used in a method of treatment of a  
 XX mammal and prevention of disorders caused by the aberrant protein  
 XX expression or activity. The polypeptides can be used as molecular weight  
 XX markers, food supplements, and in antibody production. The polypeptides  
 XX are used to identify compounds which bind to the polypeptides.  
 XX Polynucleotides of the invention are used as probes and primers, for  
 XX sequencing, for chromosome or gene mapping, in the production of  
 XX recombinant proteins, and in generating anti-sense DNA or RNA and in gene  
 XX therapy. Polypeptides of the invention can be used to target drugs to a  
 XX tumour, in assays to determine biological activity, to raise  
 XX antibodies/elicite an immune response, to determine quantitative  
 XX levels, as tissue markers, and to isolate receptors or ligands.  
 XX Polypeptides of the invention may also be useful in treating platelet  
 XX disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
 XX ligament and/or nerve tissue, wound healing, treating burns, promoting  
 XX the proliferation, differentiation and survival of stem cells, as a  
 XX contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
 XX Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
 XX sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
 XX fungal infection or from autoimmunity, cancer, allergy, asthma, graft-  
 XX versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory  
 XX diseases, nervous system disorders, and infection. The present sequence  
 XX represents a protein of the invention.

XX Sequence 1188 AA;  
 SQ  
 Query Match 77.9%; Score 926; DB 4; Length 1188;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1126; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 MDLPGLVAVWALSILWPGFTDTFNMDTRKPRVIGPSRTAFPGYTVQOHDISGNKWLWVGA 60  
 Db 1 MDLPGLVAVWALSILWPGFTDTFNMDTRKPRVIGPSRTAFPGYTVQOHDISGNKWLWVGA 60  
 Qy 61 PLETNGYQKTGDVYKCPVIHGNCTKLNLRVTLNVSERKDNMRGLSLATNPKDNSFLA 120  
 Db 61 PLETNGYQKTGDVYKCPVIHGNCTKLNLRVTLNVSERKDNMRGLSLATNPKDNSFLA 120  
 Qy 121 CSPLSHSGSSSYTTCGKSVNSNFRSKTVAPALQRCQTTMDIIVLDGNSNYPWVZ 180  
 Db 121 CSPLSHSGSSSYTTCGKSVNSNFRSKTVAPALQRCQTTMDIIVLDGNSNYPWVZ 180  
 Qy 181 VQHFLINILKXFYIGGQIQGVVGYGDDVVEHFLNDYRSVKDVEAASHIEORGGTET 240  
 Db 181 VQHFLINILKXFYIGGQIQGVVGYGDDVVEHFLNDYRSVKDVEAASHIEORGGTET 240  
 Qy 241 RTAFGIEFARSEAFQKGRKGAKVMIVITGESHSDPLEKYVIOQSERDNVTRVAVL 300  
 Db 241 RTAFGIEFARSEAFQKGRKGAKVMIVITGESHSDPLEKYVIOQSERDNVTRVAVL 300  
 Qy 301 GYNERGINPETFLNEIKYIASDPDDKHFFNVVDEAALKDIVDALGDRIFSLGNTKNET 360  
 Db 301 GYNERGINPETFLNEIKYIASDPDDKHFFNVVDEAALKDIVDALGDRIFSLGNTKNET 360  
 Qy 361 SPGLEMSQTFSSSHVVEDGVLGAVGWNGAVLKETSAGKVIPLRSYLFKEPPEELKN 420  
 Db 361 SPGLEMSQTFSSSHVVEDGVLGAVGWNGAVLKETSAGKVIPLRSYLFKEPPEELKN 420  
 Qy 421 HGAYLGYVTSVSSRQGRVVVAGAPRHNHKGKILPTMHNRSITLTIQAMGQOIGSYF 480  
 Db 421 HGAYLGYVTSVSSRQGRVVVAGAPRHNHKGKILPTMHNRSITLTIQAMGQOIGSYF 480  
 Qy 481 GSEITSDVDGDTVDLLVGAPMYFNGRGRKVVYVELRQNRVYVYNGTLKDSHSYQNA 540  
 Db 481 GSEITSDVDGDTVDLLVGAPMYFNGRGRKVVYVELRQNRVYVYNGTLKDSHSYQNA 540  
 Qy 541 RFGSSIASVRDLNQSNDVNDVVGAPLEDNHAGAIYIFHGFRGSIILKTPKORITASELATG 600  
 Db 541 RFGSSIASVRDLNQSNDVNDVVGAPLEDNHAGAIYIFHGFRGSIILKTPKORITASELATG 600  
 Qy 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILMSRPVQVQINASLHFPFSKINIFHRDCK 660  
 Db 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILMSRPVQVQINASLHFPFSKINIFHRDCK 660  
 Qy 661 RSGRDATCLAAFLCPTPIFLAPHFQTTTIGIRYNATMDERRVTPRAHLDGEGDRTNRAV 720  
 Db 661 RSGRDATCLAAFLCPTPIFLAPHFQTTTIGIRYNATMDERRVTPRAHLDGEGDRTNRAV 720  
 Qy 721 LLSQSQELCERINFHVLDTADYVKPVTFSVEYSLEDDPHGMLDDGWFTTLRSVPFWNG 780  
 Db 721 LLSQSQELCERINFHVLDTADYVKPVTFSVEYSLEDDPHGMLDDGWFTTLRSVPFWNG 780  
 Qy 781 CNEDEHCVDPDLVDARSOLPTAMEYCORVLRKPAODCSAYTILSPDTTFFIESTEQRYAV 840  
 Db 781 CNEDEHCVDPDLVDARSOLPTAMEYCORVLRKPAODCSAYTILSPDTTFFIESTEQRYAV 840  
 Qy 841 EATLENRGENAVSTVLNISQSANLQFASLIQKEDSDGSIKEDNEERLRKQVCNVSYPFF 900  
 Db 841 EATLENRGENAVSTVLNISQSANLQFASLIQKEDSDGSIKEDNEERLRKQVCNVSYPFF 900  
 Qy 901 RAKAKVAFRLDSERFSKSIPLHLELELAAGSDSNERDSTKEDNVAPLPHLYKEDVLT 960  
 Db 901 RAKAKVAFRLDSERFSKSIPLHLELELAAGSDSNERDSTKEDNVAPLPHLYKEDVLT 960  
 Qy 961 RSSLSHYEVLKSSLRDYGIGPPFSCIPRIQNLGLFPIHGMKTIPIATRSNRL 1020

```

Db 961 RSSLSHVEYKPNSSLERYDGGPPFSCIPRIQNLGLFPFIHGMWKITPIATRSNRL 1020
Qy 1021 KLRLDLTDEANTSCNIGNSTEYRTPVVEEDLRRAPQLNHSNSDVVSINCMRLVFNQEI 1080
Db 1021 KLRLDLTDEANTSCNIGNSTEYRTPVVEEDLRRAPQLNHSNSDVVSINCMRLVFNQEI 1080
Qy 1081 NFHLIGNLWLSKALKYKMKIMVNAALQROFHSPFIFREDDPSRQI 1128
Db 1081 NFHLIGNLWLSKALKYKMKIMVNAALQROFHSPFIFREDDPSRQI 1128

RESULT 4
AAB50085
ID AAB50085 standard; protein; 1188 AA.
AC AAB50085;
XX
DT 19-MAR-2001 (first entry)
DE Human A259.
XX
KW Human; A259; integrin alpha subunit; fibrosis; liver disease; lung;
KW kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX

FH Key Location/Qualifiers
FT Domain 1..1141
FT Peptide 1..22 /label= Extracellular_domain
FT Protein 23..1188 /label= Signal_peptide
FT Domain 39..74 /label= Mature_protein
FT Domain /label= Integrin_alphasubunit_repeat_domain_#1
FT Domain 115..157 /label= Integrin_alphasubunit_repeat_domain_#2
FT Domain 164..345 /label= I_domain
FT Domain 367..392 /label= I_domain
FT Domain 421..455 /label= Integrin_alphasubunit_repeat_domain_#3
FT Domain 478..516 /label= Integrin_alphasubunit_repeat_domain_#4
FT Domain 540..575 /label= Integrin_alphasubunit_repeat_domain_#5
FT Domain 602..640 /label= Integrin_alphasubunit_repeat_domain_#6
FT Domain 1142..1164 /label= Integrin_alphasubunit_repeat_domain_#7
FT Domain 1165..1188 /label= Transmembrane_domain
FT Domain /label= Cytoplasmic_domain

XX WO200073339-A1.
XX
XX 07-DEC-2000.
XX
XX 15-MAY-2000; 2000WO-US013262.
XX
XX 28-MAY-1999; 99US-00322790.
XX
XX 27-APR-2000; 2000US-00561263.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Pan Y, Lora JW;
XX
XX WPI; 2001-041142/05.
XX
XX N-PSDB; AAC91901, AAC91902.
XX
XX Nucleic acid encoding alpha-integrin subunits, useful for treatment and
XX diagnosis of fibrosis, e.g. of the liver.

```

```

XX Claim 8; Fig 1; 164pp; English.
PS
CC The present sequence is human integrin alpha subunit. A259. A259 is
CC homologous with the alpha1 and alpha10 integrin subunits and is
CC overexpressed in fibrosis. A259 is implicated in regulation of
CC proliferation, differentiation and/or function of many different cell
CC types. Inhibitors of A259 activity are useful for the treatment of liver
CC disease, particularly fibrosis, and also fibrosis in other organs
CC (specifically lung and kidney). In addition, A259 can be used for
CC treatment and prevention of cancer, osteoporosis, acute myeloid
CC leukaemia, HIV infection, and rheumatoid arthritis
XX
XX Sequence 1188 AA;
XX
Query Match 77.9%; Score 926; DB 4; Length 1188;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1126; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MDLPRGLVAVWALSINPGFTDTFNMDFKPRVPGSRTAPRGYVQOHDISGNKWLAVGA 60
Db 1 MDLPRGLVAVWALSINPGFTDTFNMDFKPRVPGSRTAPRGYVQOHDISGNKWLAVGA 60
Qy 61 PLETNQYQKTGDVYKCPVIHGNCTKLNGLRVTLNSVSRKDNMRLGLSLATNPKDNSFLA 120
Db 61 PLETNQYQKTGDVYKCPVIHGNCTKLNGLRVTLNSVSRKDNMRLGLSLATNPKDNSFLA 120
Qy 121 CSPLMSHECGSSYTTGMCSEKRVNSFRPSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
Db 121 CSPLMSHECGSSYTTGMCSEKRVNSFRPSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
Qy 181 VQHFILINILKKFYIGPGQIQGVVVOYGEDVVEHFLNDYRSVKDVAASHIEORGCTET 240
Db 181 VQHFILINILKKFYIGPGQIQGVVVOYGEDVVEHFLNDYRSVKDVAASHIEORGCTET 240
Qy 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDSEHSDPDLKVKIQSERDMVTVYAVAVL 300
Db 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDSEHSDPDLKVKIQSERDMVTVYAVAVL 300
Qy 301 GYVNRGINPETFLNEIKYIASDDDDKHFFNVDTDAALKDVIDALGDRIFSELTGNKET 360
Db 301 GYVNRGINPETFLNEIKYIASDDDDKHFFNVDTDAALKDVIDALGDRIFSELTGNKET 360
Qy 361 SFGLEMSQTGFSSHVVEDGVLGAVGAYDMNGAVLKETSAGKVIPLRESYLKPFPEELKN 420
Db 361 SFGLEMSQTGFSSHVVEDGVLGAVGAYDMNGAVLKETSAGKVIPLRESYLKPFPEELKN 420
Qy 421 HGAYLGTVTVSVSSRQGRVYVAGAPRNTGKVIPLTMMNRSLTIHQAMRGQOIGSYF 480
Db 421 HGAYLGTVTVSVSSRQGRVYVAGAPRNTGKVIPLTMMNRSLTIHQAMRGQOIGSYF 480
Qy 481 GSEITSDIDGDGVTDLVLLVGAPMYFNEGRERKGVYVYELRQNFVYNGTLKDSHSYQNA 540
Db 481 GSEITSDIDGDGVTDLVLLVGAPMYFNEGRERKGVYVYELRQNFVYNGTLKDSHSYQNA 540
Qy 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNAGAIYIFHGFRGSLKTPKORITASELATG 600
Db 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNAGAIYIFHGFRGSLKTPKORITASELATG 600
Qy 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILMSRPVQVINALSHLFEPSKINIFHRDCK 660
Db 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILMSRPVQVINALSHLFEPSKINIFHRDCK 660
Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTVGIRYNATMDERRYTPRAHLDEGDRFTTRAV 720
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTVGIRYNATMDERRYTPRAHLDEGDRFTTRAV 720
Qy 721 LLSSGQELCERINFHVLDTDYVKPVTFSVEYSLEDDHGMPLDDGPTTLRVSVFPWNG 780
Db 721 LLSSGQELCERINFHVLDTDYVKPVTFSVEYSLEDDHGMPLDDGPTTLRVSVFPWNG 780
Qy 781 CNEDEHCVDPDLVDARSDELPTAMEYCORVLRKPAQDCSAYTLSPDTTVFIESTQRVAV 840

```

|          |          |   |   |
|----------|----------|---|---|
| Db       | 781      | CNEDEHCVDPVLVLDASDLPTAMEYCORVLRKPAQDCSAYLTSPDTTTFVFIESTRQVAV  | 840   |
| Qy       | 841      | EATLENGENAYSTVLNLSQSANLQFASLIQKEDSDGSIQVNEERLQKQVCNVSYPFF   | 900   |
| Db       | 841      | EATLENGENAYSTVLNLSQSANLQFASLIQKEDSDGSIQVNEERLQKQVCNVSYPFF   | 900   |
| Qy       | 901      | RAKAKVAFRLDSFBSKIFLHLLEIELAAGSDGNERDSTKEDNVAPLRFLKYEADVLT   | 960   |
| Db       | 901      | RAKAKVAFRLDFBSKIFLHLLEIELAAGSDGNERDSTKEDNVAPLRFLKYEADVLT  | 960   |
| Qy       | 961      | RSSSLSHYEVKLNLSLRYGIGPPSCIFRIONLGLFPIHGMKITTIPATRSNRL   | 1020  |
| Db       | 961      | RSSSLSHYEVKLNLSLRYGIGPPSCIFRIONLGLFPIHGMKITTIPATRSNRL   | 1020  |
| Qy       | 1021     | KLKDFLTDEANTSCNIGNSTEVRPTVEEDLRAAPOLNHSNDVSVINCLVNPQEI  | 1080  |
| Db       | 1021     | KLKDFLTDEANTSCNIGNSTEVRPTVEEDLRAAPOLNHSNDVSVINCLVNPQEI  | 1080  |
| Qy       | 1081     | NFHLGLNMLRSKALKYKSKIMVNAALQRFHSPFIFREBPSRQI   | 1128  |
| Db       | 1081     | NFHLGLNMLRSKALKYKSKIMVNAALQRFHSPFIFREBPSRQI   | 1128  |
| RESULT 5 |          |   |   |
| AAU10551 |          |   |   |
| ID       | AAU10551 | standard; protein; 1188 AA.   |   |
| XX       | AC       | AAU10551;   |   |
| XX       | DT       | 14-FEB-2002 (first entry)   |   |
| XX       | DB       | Human A259 polypeptide.   |   |
| XX       | KW       | Human; A259; integrin alpha subunit; integrin alpha 10; secreted protein; liver disease; fibrosis; lung; kidney; bone associated disorder; blood; cartilage associated disorder; haematopoietic disorder; bone marrow; immune related disease; apoptotic disorder; neuronal tissue disease; neurodegenerative disease; gene therapy; cancer; cyostatic; osteopathic; nephrotropic; immunomodulator; anti-inflammatory; neuroprotective; antiarthritic; antianemic; antiallergic; antiasthmatic; dermatological; antidiabetic; anticonvulsant; antiparkinsonian. |   |
| XX       | OS       | Homo sapiens.   |   |
| XX       | PH       | Key   | Location/Qualifiers   |
| FT       | Domain   | 1. .1141  | /note= "Extracellular domain"   |
| FT       | Peptide  | 1. .22  | /note= "Signal peptide"   |
| FT       | Protein  | 23. .1188   | /note= "Mature human A259"  |
| FT       | Domain   | 37. .90   | /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549" |
| FT       | Domain   | 115. .157   | /note= "Integrin alpha repeat domain"   |
| FT       | Domain   | 164. .345   | /note= "I domain or Von Willebrand Factor type A domain"  |
| FT       | Domain   | 367. .392   | /note= "Integrin alpha repeat domain"   |
| FT       | Domain   | 421. .472   | /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549" |
| FT       | Domain   | 476. .532   | /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549" |
| FT       | Domain   | 538. .593   | /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549" |
| FT       | Domain   | 600. .654   | /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549" |
| FT       | Domain   | 1142. .1164   | /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549" |

|    |   |  |     |
|----|---|--|-----|
| FT | Domain  | /note= "Transmembrane domain"                                |     |
| FT | 1165. .1188   |  |     |
| XX | PN  | /note= "Cytoplasmic domain"                                  |     |
| XX | WO200181414-A2.   |  |     |
| XX | 01-NOV-2001.  |  |     |
| XX | 27-APR-2001; 2001WO-US013536.   |  |     |
| XX | 27-APR-2000; 2000US-00561263.   |  |     |
| XX | (MILL-)   | MILLENNIUM PHARM INC.  |     |
| XX | Pan Y, Lora J;  |  |     |
| XX | WPI; 2002-0411397/05.   |  |     |
| DR | N-PSDB; AAS16873.   |  |     |
| XX | New A259 nucleic acids and polypeptides, which comprise integrin alpha subunit, useful for diagnosing, preventing or treating e.g. liver disease, kidney or lung fibrosis, cancers, blood disorders or immune related diseases.   |  |     |
| XX | Claim 9; Fig 1; 168pp; English.   |  |     |
| XX | The invention relates to human and murine A259 nucleic acid molecules which encode secreted proteins with homology to integrin alpha subunits, specifically to integrin alpha 10. The A259 polypeptide and nucleic acid are useful for treating liver disease or fibrosis particularly kidney fibrosis or lung fibrosis. The A259 polypeptide and nucleic acid are also useful for diagnosing, preventing or treating cartilage and bone associated disorders (such as bone cancer, achondroplasia, myeloma, fibrous dysplasia, scoliosis, osteoarthritis, osteosarcoma and acute myeloid leukaemia, haemophilia, anaemia and thalassemia), immune related diseases (such as HIV, viral infections, cancers, T cell autoimmune disorders e.g. AIDS, and allergic inflammatory disorders e.g. asthma and psoriasis), apoptotic disorders (such as systemic lupus erythematosus and insulin-dependent diabetes mellitus), diseases of the neuronal tissues (such as epilepsy and muscular dystrophy) and neurodegenerative diseases (such as Parkinson's disease and Huntington's disease). This sequence represents the human A259 polypeptide |  |     |
| XX | Sequence 1188 AA;   |  |     |
| Qy | Query Match   | 77.9%; Score 926; DB 5; Length 1188;                         |     |
| Db | Best Local Similarity   | 99.8%; Pred. No. 0;  |     |
| Qy | Matches 1126; Conservative  | 0; Mismatches 2; Indels 0; Gaps 0;                           |     |
| Qy | 1   | MDLPEGLVAVNALSLWPGFTTETNMDTKPRVPGSRTAFPGYTVQOHDISGNKLVVGA    | 60  |
| Db | 1   | MDLPEGLVAVNALSLWPGFTTETNMDTKPRVPGSRTAFPGYTVQOHDISGNKLVVGA    | 60  |
| Qy | 61  | PLETNGYQKTGDVYKCPVIHGNCTKLNLRVTLSNVSRKDNMRGLSLATNPKDNSFLA    | 120 |
| Db | 61  | PLETNGYQKTGDVYKCPVIHGNCTKLNLRVTLSNVSRKDNMRGLSLATNPKDNSFLA    | 120 |
| Qy | 121   | CSPLWSHCGSSYYTTCMSRVNFRFSKTVAPALQRCOTYMDIVLDGNSLYPWVE        | 180 |
| Db | 121   | CSPLWSHCGSSYYTTCMSRVNFRFSKTVAPALQRCOTYMDIVLDGNSLYPWVE        | 180 |
| Qy | 181   | VQHFILNLIKFFIYIGPGQIQGVVYQYGEDVVHFEHLNDYRSVKDVVEAASHIEQGGTET | 240 |
| Db | 181   | VQHFILNLIKFFIYIGPGQIQGVVYQYGEDVVHFEHLNDYRSVKDVVEAASHIEQGGTET | 240 |
| Qy | 241   | RTAFGIEFARSEAFQKGRKGAKKVMIVITDGEHSDSPLEKVIQSSRDVNVRYAVAVL    | 300 |
| Db | 241   | RTAFGIEFARSEAFQKGRKGAKKVMIVITDGEHSDSPLEKVIQSSRDVNVRYAVAVL    | 300 |
| Qy | 301   | GYNNRGINPETFLNEIKYIASDPDDKHFFNVVTDEAALKDVIDALGDRIFSLGNTKNET  | 360 |
| Db | 301   | GYNNRGINPETFLNEIKYIASDPDDKHFFNVVTDEAALKDVIDALGDRIFSLGNTKNET  | 360 |

QY 361 SFGLEMSQTGFSSHVHVDGVLGAYDNGAVLKETSGAGVILPRLSYLKEPPEELKN 420  
 DB |||||  
 QY 361 SFGLEMSQTGFSSHVHVDGVLGAYDNGAVLKETSGAGVILPRLSYLKEPPEELKN 420  
 DB |||||  
 QY 421 HGAYLGVTYVTSVVSQRQGVYVAGAPRPNHETGKVLFTMHNNRSLTIHOAMRGQOIGSYF 480  
 DB |||||  
 QY 421 HGAYLGVTYVTSVVSQRQGVYVAGAPRPNHETGKVLFTMHNNRSLTIHOAMRGQOIGSYF 480  
 DB |||||  
 QY 481 GSEITSVDIDGQVTDVLLVGAPMYENEGREGKVVYVELRQNRVFNCTLKDSYQNA 540  
 DB |||||  
 QY 481 GSEITSVDIDGQVTDVLLVGAPMYENEGREGKVVYVELRQNRVFNCTLKDSYQNA 540  
 DB |||||  
 QY 541 RFGSSIASVRDLNODSYNDVVVYGAPLEDNHAGAIYIFHGFGRSILKTPKQRITASELATG 600  
 DB |||||  
 QY 541 RFGSSIASVRDLNODSYNDVVVYGAPLEDNHAGAIYIFHGFGRSILKTPKQRITASELATG 600  
 DB |||||  
 QY 601 LOYFGCSIHGOLDIADNGLDANGALGNVILWSPVVOINASIHFPSPKINIFHRDCK 660  
 DB |||||  
 QY 601 LOYFGCSIHGOLDIADNGLDANGALGNVILWSPVVOINASIHFPSPKINIFHRDCK 660  
 DB |||||  
 QY 661 RSGRATCLAAFLCTPIFLAPHFQTTVIGIRYNATMDERRYTTPRAHLDEGDRTRNAV 720  
 DB |||||  
 QY 661 RSGRATCLAAFLCTPIFLAPHFQTTVIGIRYNATMDERRYTTPRAHLDEGDRTRNAV 720  
 DB |||||  
 QY 721 LLSGQELCERINFHVLDTADYVKVPTFSVEYSLEDPOHGMPLDDGWPTTLRVSPFWNG 780  
 DB |||||  
 QY 721 LLSGQELCERINFHVLDTADYVKVPTFSVEYSLEDPOHGMPLDDGWPTTLRVSPFWNG 780  
 DB |||||  
 QY 781 CNEDEHCVPLVDLARSDLPTAMEYCORVLRPAQDCSAYTUSFDTTFIETSTQRVAV 840  
 DB |||||  
 QY 781 CNEDEHCVPLVDLARSDLPTAMEYCORVLRPAQDCSAYTUSFDTTFIETSTQRVAV 840  
 DB |||||  
 QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSTEVCNEBRRLQKQVCNVSYPFF 900  
 DB |||||  
 QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSTEVCNEBRRLQKQVCNVSYPFF 900  
 DB |||||  
 QY 901 RAKAVAFELDFEFSKIFLHLELELAGSDNSNERDSTKEDNVAPLPHLYKAEADVLPT 960  
 DB |||||  
 QY 901 RAKAVAFELDFEFSKIFLHLELELAGSDNSNERDSTKEDNVAPLPHLYKAEADVLPT 960  
 DB |||||  
 QY 961 RSSLSHVEVKLNSSLERYDGIGPPFCIFRIONLGLPPIHGXMMKIPIPIATRSNRL 1020  
 DB |||||  
 QY 961 RSSLSHVEVKLNSSLERYDGIGPPFCIFRIONLGLPPIHGXMMKIPIPIATRSNRL 1020  
 DB |||||  
 QY 1021 KLDELDTDEANTSCINWNGSTERYPTVEEDLRAPQLNHSNDVVSINCNIRLVPNOEI 1080  
 DB |||||  
 QY 1021 KLDELDTDEANTSCINWNGSTERYPTVEEDLRAPQLNHSNDVVSINCNIRLVPNOEI 1080  
 DB |||||  
 QY 1081 NFHLLGNLWRLSLKALKYKSMIMVNAALQROFHSPPFIREDPSROI 1128  
 DB |||||  
 QY 1081 NFHLLGNLWRLSLKALKYKSMIMVNAALQROFHSPPFIREDPSROI 1128  
 DB |||||  
 RESULT 6  
 AAU14467  
 ID AAU14467 standard; protein; 1188 AA.  
 AC AAU14467;  
 XX AAU14467;  
 XX AAU14467;  
 DT 24-OCT-2001 (first entry)  
 XX Human novel protein #338.  
 DE Human; novel protein; Antianaemic; osteopathic; antiinflammatory;  
 KW immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic;  
 KW anticonvulsant; antirheumatic; cerebroprotective; antifungal; antiviral;  
 KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;  
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
 KW tissue regeneration; immune disorder.  
 XX Homo sapiens.  
 OS

XX WO200155437-A2.  
 PN  
 XX  
 PD 02-AUG-2001.  
 XX  
 XX 25-JAN-2001; 2001WO-US002623.  
 PP  
 XX  
 PR 25-JAN-2000; 2000US-00491404.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX  
 XX Tang YT, Liu C, Drmanac RT;  
 PI  
 XX  
 DR WPI; 2001-451939/48.  
 DR N-PSDB; AAS22772.  
 XX  
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,  
 PT nervous system disorders, and for regenerating bone and cartilage.  
 XX  
 PS Example 4; Page 828-831; 894pp; English.  
 XX  
 CC The invention relates to polynucleotides encoding novel human proteins or  
 CC their active domains. The polypeptides, polynucleotides and antibodies  
 CC raised against the polypeptides are used in a method of treatment of a  
 CC mammal and prevention of disorders caused by the aberrant protein  
 CC expression or activity. The polypeptides can be used as molecular weight  
 CC markers, food supplements, and in antibody production. The polypeptides  
 CC are used to identify compounds which bind to the polypeptides.  
 CC Polynucleotides of the invention are used as probes and primers, for  
 CC sequencing, for chromosome or gene mapping, in the production of  
 CC recombinant proteins, and in generating anti-sense DNA or RNA and in gene  
 CC therapy. Polypeptides of the invention can be used to target drugs to a  
 CC tumour, in assays to determine biological activity, to raise  
 CC antibodies/elicite an immune response, to determine quantitative protein  
 CC levels, as tissue markers, and to isolate receptors or ligands.  
 CC Polypeptides of the invention may also be useful in treating platelet  
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
 CC the proliferation, differentiation and survival of stem cells, as a  
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia, as a  
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
 CC fungal infection or from autoimmunity, cancer, allergy, asthma, graft-  
 CC versus-host disease, eczema, haemophilia, thrombosis, anti-inflamatory  
 CC diseases, nervous system disorders, and infection. The present sequence  
 CC represents a protein of the invention  
 XX  
 SQ Sequence 1188 AA;

Query Match 76.7%; Score 911; DB 4; Length 1188;  
 Best Local Similarity 100.0%; Pred.No. 0;  
 Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDLPRLGVAVWALSINWEGFTDTFNMOTRKPVRIPGSRTPAFGYTVQQHDSGNKWLVVGA 60  
 DB 1 MDLPRLGVAVWALSINWEGFTDTFNMOTRKPVRIPGSRTPAFGYTVQQHDSGNKWLVVGA 60  
 QY 61 PLETNGYQKTDVYKCPVHGNCCTKMLGRVTLNVSERKDNMLRGLSLATNPKDNGSFLA 120  
 DB 61 PLETNGYQKTDVYKCPVHGNCCTKMLGRVTLNVSERKDNMLRGLSLATNPKDNGSFLA 120  
 QY 121 CSLPLWSHECGSSYYTTCMCSRVSNSNFPFSKTVAPALORCQTYMDIVLVDGSSNIYVWE 180  
 DB 121 CSLPLWSHECGSSYYTTCMCSRVSNSNFPFSKTVAPALORCQTYMDIVLVDGSSNIYVWE 180  
 QY 181 VQHEFLINILKKFYIGPGQIQGVVYQGEDVVHBFHLNDYRSVKDVAASHIEORGGTET 240  
 DB 181 VQHEFLINILKKFYIGPGQIQGVVYQGEDVVHBFHLNDYRSVKDVAASHIEORGGTET 240  
 QY 241 RTAFGIFPAREAPQKGRKGAKKVMITVITDGHSDSPDLSEKVIQOSERDNVTYAVAVL 300  
 DB 241 RTAFGIFPAREAPQKGRKGAKKVMITVITDGHSDSPDLSEKVIQOSERDNVTYAVAVL 300



301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVVTDEAALKDVIDALGDRIFSLEGTNKNET 360  
301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVVTDEAALKDVIDALGDRIFSLEGTNKNET 360  
361 SFGLEMSQTGFSSHHVDEGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPBELKN 420  
361 SFGLEMSQTGFSSHHVDEGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPBELKN 420  
421 HGAYLGYTVTSVSSRQGRVYVAGAPFRFNTGKVLFTMHNNRSLTIHOAMRGQQIGSYF 480  
421 HGAYLGYTVTSVSSRQGRVYVAGAPFRFNTGKVLFTMHNNRSLTIHOAMRGQQIGSYF 480  
481 GSEITSDVDGVDGVTDLVLLGAPMYFNEGREKGVYVELRQNRVYVINGTLKDSHSYQNA 540  
481 GSEITSDVDGVDGVTDLVLLGAPMYFNEGREKGVYVELRQNRVYVINGTLKDSHSYQNA 540  
541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFGRGSLTKPQRITASELATG 600  
541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFGRGSLTKPQRITASELATG 600  
601 LOYFGCSIHQDLNEDGLDLAVGALGNVILKSRPVVQINASLHFPESKINIFHRDCK 660  
601 LOYFGCSIHQDLNEDGLDLAVGALGNVILKSRPVVQINASLHFPESKINIFHRDCK 660  
661 RSGRDATCLAAFLCFTPIFLAPHFQTTVYVIRYNATMDERRYTTPRAHLDGGDRFTNRV 720  
661 RSGRDATCLAAFLCFTPIFLAPHFQTTVYVIRYNATMDERRYTTPRAHLDGGDRFTNRV 720  
721 LLSGGQELCERINFHVLDTADYVKVPTFSVEYSLEDDPDHGMELDDGWPTTLRVSVFPMNG 780  
721 LLSGGQELCERINFHVLDTADYVKVPTFSVEYSLEDDPDHGMELDDGWPTTLRVSVFPMNG 780  
781 CNEDEHCVPLVLDARSDLSTANBYCORVLRKPAQDCSAYTILSFDTTVFIIESTRORVAV 840  
781 CNEDEHCVPLVLDARSDLSTANBYCORVLRKPAQDCSAYTILSFDTTVFIIESTRORVAV 840  
841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900  
841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900  
901 RAKAKVAFRLD 911  
901 RAKAKVAFRLD 911

RESULT 7  
ADE09956  
ID ADE09956 standard; protein; 1188 AA.  
XX AC ADE09956;  
XX DT 29-JAN-2004 (first entry)  
XX DE Novel protein-related contig polypeptide sequence #544.  
XX KW novel gene; novel protein; tissue marker; molecular weight marker;  
XX KM chromosome marker; genetic disorder; contig.  
XX OS Unidentified.  
XX PW W02003054152-A2.  
XX PD 03-JUL-2003.  
XX PF 10-DEC-2002; 2002W0-US039555.  
XX PR 10-DEC-2001; 2001US-0339739P.  
XX PR 11-DEC-2001; 2001US-0339453P.  
XX PR 14-MAR-2002; 2002US-0365091P.  
XX PR 14-MAR-2002; 2002US-0365384P.  
XX PR 12-APR-2002; 2002US-0372615P.  
XX PR 22-APR-2002; 2002US-00128558.

24-APR-2002; 2002US-0376045P.  
(HYSE-) HYSEQ INC.  
Tang YT, Asundi V, Goodrich RN, Ren P, Zhang J, Zhao QA, Wang J;  
Ghosh M, Xue AJ, Wehrman I, Weng G, Zhou P, Drmanac RT, Wang Z;  
Ma Y, Wang D, Chen R, Xu C, Boyle BU;  
WPI; 2003-569235/53.  
New polynucleotides, useful for expressing recombinant proteins for  
analysis, characterization or therapeutic use, or as markers for tissues  
in which the corresponding protein is preferentially expressed.  
Disclosure; SEQ ID NO 3022; 1177bp; English.  
The invention comprises the amino acid and coding sequences of novel  
proteins. The DNA and protein sequences of the invention are useful as:  
markers for tissues in which the corresponding protein is preferentially  
expressed; as molecular weight markers on gels, as chromosome markers or  
tags; to identify chromosomes or to map related gene positions; and to  
compare with endogenous DNA sequences in patients to identify potential  
genetic disorders. The present amino acid sequence was used in the  
exemplification of the invention.  
Sequence 1188 AA;  
Query Match 76.7%; Score 911; DB 7; Length 1188;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDLPRLGLVAVALSLLWPGFTDTFNNMTRKPRVPGSRATFFGYTVQQHDSGNKWLWVGA 60  
DB 1 MDLPRLGLVAVALSLLWPGFTDTFNNMTRKPRVPGSRATFFGYTVQQHDSGNKWLWVGA 60  
QY 61 PLETTYGKTDGVYKCPVIHGNCTKLNLRVTLSNVSRKDNMRGLSLATPKDNSFLA 120  
DB 61 PLETTYGKTDGVYKCPVIHGNCTKLNLRVTLSNVSRKDNMRGLSLATPKDNSFLA 120  
QY 121 CSPLWSHSCGSSYYTTGMCNRVNSFRSKTVAPALQRCQTYMDIVIVLDGNSNTPWVE 180  
DB 121 CSPLWSHSCGSSYYTTGMCNRVNSFRSKTVAPALQRCQTYMDIVIVLDGNSNTPWVE 180  
QY 181 VOHFLINILKRPYIGPGQIQGVGVQYGVGVVHPEHNDYRSVKDVAAASHIEQGGTET 240  
DB 181 VOHFLINILKRPYIGPGQIQGVGVQYGVGVVHPEHNDYRSVKDVAAASHIEQGGTET 240  
QY 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITGESHSDSPDLEKVIQOQSRDNVTRVAVL 300  
DB 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITGESHSDSPDLEKVIQOQSRDNVTRVAVL 300  
QY 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVVTDEAALKDVIDALGDRIFSLEGTNKNET 360  
DB 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVVTDEAALKDVIDALGDRIFSLEGTNKNET 360  
QY 361 SFGLEMSQTGFSSHHVDEGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPBELKN 420  
DB 361 SFGLEMSQTGFSSHHVDEGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPBELKN 420  
QY 421 HGAYLGYTVTSVSSRQGRVYVAGAPFRFNTGKVLFTMHNNRSLTIHOAMRGQQIGSYF 480  
DB 421 HGAYLGYTVTSVSSRQGRVYVAGAPFRFNTGKVLFTMHNNRSLTIHOAMRGQQIGSYF 480  
QY 481 GSEITSDVDGVDGVTDLVLLGAPMYFNEGREKGVYVELRQNRVYVINGTLKDSHSYQNA 540  
DB 481 GSEITSDVDGVDGVTDLVLLGAPMYFNEGREKGVYVELRQNRVYVINGTLKDSHSYQNA 540  
QY 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFGRGSLTKPQRITASELATG 600  
DB 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFGRGSLTKPQRITASELATG 600  
QY 601 LOYFGCSIHQDLNEDGLDLAVGALGNVILKSRPVVQINASLHFPESKINIFHRDCK 660  
DB 601 LOYFGCSIHQDLNEDGLDLAVGALGNVILKSRPVVQINASLHFPESKINIFHRDCK 660



Db 601 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILSRPVVQINASLHFPSPKINIPHRDCK 660  
QY 661 RSGRDATCLAAFLCTPIPLADPHFQTTVGIRYNATMBERRYTPRAHLDEGGDRFTNRV 720  
Db 661 RSGRDATCLAAFLCTPIPLADPHFQTTVGIRYNATMBERRYTPRAHLDEGGDRFTNRV 720  
QY 721 LLSGQELCERINFHLVDADYVVKVTSVEYSLEDDPHGPMLDGWPFTLRVSVPPFWNG 780  
Db 721 LLSGQELCERINFHLVDADYVVKVTSVEYSLEDDPHGPMLDGWPFTLRVSVPPFWNG 780  
QY 781 CNEDEHCVDPDLVDARSDLPTAMEYCORVLRKPAODCSAYTLSPDTTPILESFTRQVAV 840  
Db 781 CNEDEHCVDPDLVDARSDLPTAMEYCORVLRKPAODCSAYTLSPDTTPILESFTRQVAV 840  
QY 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGSEICVNEERRLQKQVCNVSYPFF 900  
Db 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGSEICVNEERRLQKQVCNVSYPFF 900  
QY 901 RAKAKVAFRLD 911  
Db 901 RAKAKVAFRLD 911

RESULT 8  
ID AAB25582  
XX AAB25582 standard; protein; 1189 AA.  
AC AAB25582;  
DT 21-NOV-2000 (first entry)  
XX ITGAL1 protein encoded by human secreted protein gene #7.  
DE Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;  
KW antirheumatic, dermatological; antiproliferative; antiarteriosclerotic;  
KW anticancer; vulnary; antiviral; antibacterial; antifungal;  
KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;  
KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;  
KW Crohn's disease; nephritis; hyperproliferative disorder;  
KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;  
KW melanoma; lymphoma; wound healing; human; chromosome 15 q22.3-23.  
XX Homo sapiens.  
OS  
XX WO200029435-A1.  
XX 25-MAY-2000.  
XX 27-OCT-1999; 99WO-US025031.  
XX 28-OCT-1998; 98US-0105971P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;  
PI Greene JM;  
XX WPI; 2000-387742/33.  
XX N-PSDB; AAB80612.  
XX Isolated nucleic acid molecules encoding human secreted proteins are used  
PT for the prevention, amelioration and treatment of autoimmune,  
PT inflammatory, hyperproliferative and cardiovascular disorders, cancer,  
PT wounds, and infectious diseases.  
XX Claim 1; Fig 19A-F; 803pp; English.  
XX The present invention relates to 12 secreted human proteins and the  
CC nucleotide sequences encoding them. The polynucleotide sequences given in  
CC AAB80606-A80623 encode the 12 secreted protein sequences given in  
CC AAB25576-B25593. The human secreted proteins have various activities  
CC dependent on the tissues in which they are expressed. Examples of the  
CC activities of the proteins include: immunosuppressant; anti-inflammatory;

antiarthritic; antirheumatic, dermatological; antiproliferative;  
antiarteriosclerotic; anticancer; vulnary; antiviral; antibacterial;  
and antifungal activity. The proteins, polypeptides, agonists and  
antagonists may be used to treat prevent and/or diagnose various disease,  
disorders and conditions examples of which include: immune disorders e.g.  
Addison's disease, rheumatoid arthritis, dermatitis, and multiple  
sclerosis; inflammatory disorders e.g. inflammatory bowel disease,  
Crohn's disease and nephritis; hyperproliferative disorders such as  
paraproteinemia and purpura; cardiovascular disorders e.g. coronary  
arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The  
proteins and polynucleotide sequences may also be used in wound healing  
and the treatment of infectious diseases. The human secreted protein gene  
#7 and protein sequences are represented in sequences AAB80612 and  
AAB25582. Secreted protein gene #7 is located at position chromosome 15  
q22.3-23. Sequences AAB80652-A80661 represent genes which are related to  
the secreted protein gene#7  
XX  
SQ Sequence 1189 AA;  
Query Match 76.7%; Score 911; DB 3; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDLPRGLVAVWALSLLWPGFTDTFNDTRKPRVPGSRTAPFGYTVQOHDISGNKWLUVGA 60  
Db 1 MDLPRGLVAVWALSLLWPGFTDTFNDTRKPRVPGSRTAPFGYTVQOHDISGNKWLUVGA 60  
QY 61 PLETNQYQKTDGVYKCPVHGNCTKLNLRVTLSNVSERKDNMELGSLATNPKDNSPLA 120  
Db 61 PLETNQYQKTDGVYKCPVHGNCTKLNLRVTLSNVSERKDNMELGSLATNPKDNSPLA 120  
QY 121 CSPLWSHECCSSYYTTCMCGRVNSNFRPSKTVAPALQRCOTYNDIVILVDCGNSIYVWE 180  
Db 121 CSPLWSHECCSSYYTTCMCGRVNSNFRPSKTVAPALQRCOTYNDIVILVDCGNSIYVWE 180  
QY 181 VQFLINILKKFYIGPQIQVGVVQYGEDVVEHPLNDYRSVDVWAAASHIEQGGTET 240  
Db 181 VQFLINILKKFYIGPQIQVGVVQYGEDVVEHPLNDYRSVDVWAAASHIEQGGTET 240  
QY 241 RTAFGIEFARSEAPQKGRGAKKWMIVITDGESHSDPDLKVIQOESRDNVTYAVAVL 300  
Db 241 RTAFGIEFARSEAPQKGRGAKKWMIVITDGESHSDPDLKVIQOESRDNVTYAVAVL 300  
QY 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVYDEAALKDIDALGDRIIFSLGNTNKNET 360  
Db 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVYDEAALKDIDALGDRIIFSLGNTNKNET 360  
QY 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDNWNGAVLKETSAGKVIPLRESYLKEPPELKN 420  
Db 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDNWNGAVLKETSAGKVIPLRESYLKEPPELKN 420  
QY 421 HGAYLGVTVTSVSSRQGRVYVAGAPRFTHTGKVIPLFTMNNSLTTHQAWRGQOIGSYP 480  
Db 421 HGAYLGVTVTSVSSRQGRVYVAGAPRFTHTGKVIPLFTMNNSLTTHQAWRGQOIGSYP 480  
QY 481 GSEITSVDIDGDGVTDVLLVAGPMYFNEGRERGVYVYELRQNFVYNGTLKDSHSYQNA 540  
Db 481 GSEITSVDIDGDGVTDVLLVAGPMYFNEGRERGVYVYELRQNFVYNGTLKDSHSYQNA 540  
QY 541 RFGSSIASVADLNQDSYNDVVGAPEDNAGAIYIHPGGRGSLKTPKORITASELATG 600  
Db 541 RFGSSIASVADLNQDSYNDVVGAPEDNAGAIYIHPGGRGSLKTPKORITASELATG 600  
QY 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILSRPVVQINASLHFPSPKINIPHRDCK 660  
Db 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILSRPVVQINASLHFPSPKINIPHRDCK 660  
QY 661 RSGRDATCLAAFLCTPIPLADPHFQTTVGIRYNATMBERRYTPRAHLDEGGDRFTNRV 720  
Db 661 RSGRDATCLAAFLCTPIPLADPHFQTTVGIRYNATMBERRYTPRAHLDEGGDRFTNRV 720  
QY 721 LLSGQELCERINFHLVDADYVVKVTSVEYSLEDDPHGPMLDGWPFTLRVSVPPFWNG 780  
Db 721 LLSGQELCERINFHLVDADYVVKVTSVEYSLEDDPHGPMLDGWPFTLRVSVPPFWNG 780

Db 721 LLSGQELCERINFHVLDTADYKVPVTSVEYSLEDDHGPMLDDGHPVTLRVSVFPWNG 780  
Qy 781 CNEDEHCVPDLVLDASDLPTAMEYCORVLRKPAQDCSAYTISFDITVPIIESTRQAV 840  
Db 781 CNEDEHCVPDLVLDASDLPTAMEYCORVLRKPAQDCSAYTISFDITVPIIESTRQAV 840  
Qy 841 EATLENGENAYSTVLNISOSANLOPASLIQKSDSGSIECVNEERLQKQVNWSPYPPF 900  
Db 841 EATLENGENAYSTVLNISOSANLOPASLIQKSDSGSIECVNEERLQKQVNWSPYPPF 900  
Qy 901 RAKAKVAFRLD 911  
Db 901 RAKAKVAFRLD 911  
RESULT 9  
ABR58364  
ID ABR58364 standard; protein; 1189 AA.  
XX AC ABR58364;  
XX DT 07-JUL-2003 (first entry)  
XX DE Human NOV2a.  
XX Human; NOV; antidiabetic; anorectic; antibacterial; virucide;  
XX immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia;  
XX antiparkinsonian; antilipaeamic; gene therapy; metabolic disorder;  
XX diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;  
XX neurodegenerative disorder; Alzheimer's disease; immune disorder;  
XX haematopoietic disorder.  
XX Homo sapiens.  
XX OS  
XX PN WO2003029423-A2.  
XX PD 10-APR-2003.  
XX PF 02-OCT-2002; 2002WO-US031358.  
XX PR 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327342P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339266P.  
PR 24-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 12-APR-2002; 2002US-0371972P.  
PR 12-APR-2002; 2002US-0371980P.  
PR 17-APR-2002; 2002US-0373261P.  
PR 19-APR-2002; 2002US-0373805P.  
PR 23-APR-2002; 2002US-0374738P.  
PR 16-MAY-2002; 2002US-0381101P.  
PR 17-MAY-2002; 2002US-0381635P.  
PR 29-MAY-2002; 2002US-0383830P.  
PR 01-OCT-2002; 2002US-00262839.  
XX (CURA-) CURAGEN CORP.  
XX PA  
XX PI Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E,  
PI Edinger SR, Ellerman K, Gerlach VU, Gorman L, Guo X, Ji W,  
PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK,  
PI Rothenberg ME, Shimkets RA, Smithson G, Spytek KA, Taupier RJ,  
PI Vernet CAM, Voss EZ, Zerbussen BD, Zhong M;  
XX WPI; 2003-381625/36.

DR N-PSDB; ACC72076.  
XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or  
PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or  
PT dyslipidaemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX Claim 1; Page 105; 487pp; English.  
XX The present invention relates to novel human NOV proteins and their  
CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV  
CC proteins are useful in manufacturing a medicament for treating a syndrome  
CC associated with a human disease. The NOV proteins and coding sequences  
CC may be used to diagnose, treat or prevent metabolic disorders such as  
CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative  
CC disorders such as Alzheimer's disease or Parkinson's disease, immune  
CC disorders, haematopoietic disorders and various dyslipidaemias  
XX Sequence 1189 AA;  
SQ Query Match 76.7%; Score 911; DB 6; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDLPRGLVAVALSILWPGFTDTFNMDTKPRVIGSRTAPFGYTVQQHDISGNKWLVVGA 60  
Db 1 MDLPRGLVAVALSILWPGFTDTFNMDTKPRVIGSRTAPFGYTVQQHDISGNKWLVVGA 60  
QY 61 PLETVGYOKTGDVYKCPVIHGNCTKLNIGRVTLNVNRSERKDMRLGLSLATNPKNDSFLA 120  
Db 61 PLETVGYOKTGDVYKCPVIHGNCTKLNIGRVTLNVNRSERKDMRLGLSLATNPKNDSFLA 120  
QY 121 CSPLMSHECGSSYYTTGMSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDSNSIYPWVE 180  
Db 121 CSPLMSHECGSSYYTTGMSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDSNSIYPWVE 180  
QY 181 VOHFLINILKKFYIGPGQIQGVVQYGEDVHVEHFLNDYRSVKDVAASHIEQGGTET 240  
Db 181 VOHFLINILKKFYIGPGQIQGVVQYGEDVHVEHFLNDYRSVKDVAASHIEQGGTET 240  
QY 241 RTAFGIEFARSEAFQKGRKGAKKVMI VITDGSHESDSPLEKVIQQSERDNVTRYAVAVL 300  
Db 241 RTAFGIEFARSEAFQKGRKGAKKVMI VITDGSHESDSPLEKVIQQSERDNVTRYAVAVL 300  
QY 301 GYNNRGINPETFLNEIKYIASDDDDKHFFNVTDAAALKDIDALGDEIFSLGNTKNKET 360  
Db 301 GYNNRGINPETFLNEIKYIASDDDDKHFFNVTDAAALKDIDALGDEIFSLGNTKNKET 360  
QY 361 SFGLEMSQTFSSHVEDGVLLGAVGYDMNGAVLKETSGACKVPLRESYLKEPPEELKN 420  
Db 361 SFGLEMSQTFSSHVEDGVLLGAVGYDMNGAVLKETSGACKVPLRESYLKEPPEELKN 420  
QY 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNHTGKVLFTMHNRRSLTIHQNRGQOIGSYF 480  
Db 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNHTGKVLFTMHNRRSLTIHQNRGQOIGSYF 480  
QY 481 GSEITSDIDGDTGVTDLVLLGAPVYFNEGRERKGVYVYELQNRFPVYNGTLKDSHSYQNA 540  
Db 481 GSEITSDIDGDTGVTDLVLLGAPVYFNEGRERKGVYVYELQNRFPVYNGTLKDSHSYQNA 540  
QY 541 RFGSSIASVRDLNDSYNDVVGAPLEDNHAAGAYIPIHGFRGSIKTPKQITASELATG 600  
Db 541 RFGSSIASVRDLNDSYNDVVGAPLEDNHAAGAYIPIHGFRGSIKTPKQITASELATG 600  
QY 601 LOYFGCSIHGOLDLNEEDGLIDIAVGLGNVILSRPVVQINASHLPPSPKINIFHRDCK 660  
Db 601 LOYFGCSIHGOLDLNEEDGLIDIAVGLGNVILSRPVVQINASHLPPSPKINIFHRDCK 660  
QY 661 RSGRDATCLAAFLCFTPIFLAPHFTQTTVTVGIRYNATMDERYTTPRAHLDEGGDRPTNRAV 720  
Db 661 RSGRDATCLAAFLCFTPIFLAPHFTQTTVTVGIRYNATMDERYTTPRAHLDEGGDRPTNRAV 720  
QY 721 LLSGQELCERINFHVLDTADYKVPVTSVEYSLEDDHGPMLDDGHPVTLRVSVFPWNG 780

Db 722 LSSGQELCERINFHVLADYVYKPTFSVYSLEDPDRGMLDGGWPTLSVSPFWNG 780  
QY 781 CNEDEHCVPLVDLARSDLPTAMEYCORVLRKPAQDCSAYTSLSDTTTFVFIESTRQVAV 840  
Db 781 CNEDEHCVPLVDLARSDLPTAMEYCORVLRKPAQDCSAYTSLSDTTTFVFIESTRQVAV 840  
QY 841 EATLNRGNAYSTVLNISQSANLQFASLIQKEDSDGSGTECNERRRLOKQVNSYPPF 900  
Db 841 EATLNRGNAYSTVLNISQSANLQFASLIQKEDSDGSGTECNERRRLOKQVNSYPPF 900  
QY 901 RAKAKVAFRLD 911  
Db 901 RAKAKVAFRLD 911  
RESULT 10  
ADA27054  
ID ADA27054 standard; protein; 1189 AA.  
XX ADA27054;  
AC ADA27054;  
XX 20-NOV-2003 (first entry)  
XX Human novel secreted protein from cDNA HOBY69 #1.  
XX cytostatic; antiinflammatory; immunomodulator; neuroprotective;  
XX hemostatic; gene therapy; cancer; inflammation; immune disorder;  
XX neurological disorder; blood clotting disorder; food additive;  
XX preservative; human; secreted protein.  
XX Homo sapiens.  
XX US2003055231-A1.  
XX 20-MAR-2003.  
XX 29-OCT-2001; 2001US-00984130.  
XX 28-OCT-1998; 98US-0105971P.  
XX 27-OCT-1999; 99KO-US025031.  
XX 19-APR-2000; 2000US-0198407P.  
XX 30-OCT-2000; 2000US-0243792P.  
XX 18-APR-2001; 2001US-00836353.  
XX (NIJ/) NI J.  
XX (YOUN/) YOUNG P B.  
XX (KENN/) KENNY J J.  
XX (OLSE/) OLSEN H S.  
XX (MOOR/) MOORE P A.  
XX (WEI/) WEI Y.  
XX (GEE/) GREENE J M.  
XX (RUB/) RUBEN S M.  
XX (LIUD/) LIU D.  
XX (CROG/) CROCKER P R.  
XX Ni J, Young PB, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;  
PI Ruben SM, Liu D, Crocker PR;  
XX WPI; 2003-567103/53.  
XX N-PSDB; ADA27036.  
XX New human secreted nucleic acid molecules and polypeptides, useful for  
PT preventing, treating, or ameliorating a medical condition, such as  
PT cancer, inflammation, immune disorders, neurological and blood clotting  
PT disorders.  
XX Claim 11; Fig 19; 454pp; English.  
XX The invention relates to an isolated nucleic molecule that is at least  
CC 95% identical to 18 human cDNA sequences representing 12 novel genes  
CC encoding secreted proteins or a polynucleotide fragment of the cDNA  
CC sequence contained in American Type Culture Collection (ATCC) deposit No.

CC defined in the specification, its species homologue, a variant or allelic  
CC variant of the polynucleotide having a polynucleotide capable of  
CC hybridising under conditions the polynucleotide, where the polynucleotide  
CC does not hybridise under stringent conditions to a nucleic acid molecule  
CC having a nucleotide sequence of only A or T residues. Also included are  
CC recombinant vectors, host cells (for producing the polypeptide), the  
CC secreted polypeptide (comprising a sequence that is at least 95%  
CC identical to a polypeptide fragment, domain, epitope, full-length  
CC protein, variant, allelic variant or species homologue), antibodies that  
CC specifically bind to the polypeptides, diagnosing, treating, preventing  
CC or ameliorating a medical condition by administering the polynucleotide  
CC or the polypeptide, the gene corresponding to the cDNA sequence and  
CC identifying an activity in a biological assay (by expressing the cDNA  
CC sequence in a cell, isolating the supernatant, and detecting an activity  
CC in a biological assay and identifying the protein in the supernatant  
CC having the activity). The polypeptides, nucleic acids and antibodies are  
CC useful for diagnosing a pathological condition or a susceptibility to a  
CC pathological condition, for preventing, treating, or ameliorating a  
CC medical condition, such as cancer, inflammation and other immune  
CC disorders, neurological and blood clotting disorders (many examples are  
CC given in the specification). The nucleic acids are also useful for  
CC chromosome identification, radiation hybrid mapping or long-range  
CC restriction mapping. The polypeptides and antibodies are useful for  
CC providing immunological probes for differential identification of the  
CC tissues immunohistochemistry assays. The polypeptide, polynucleotide,  
CC agonist or antagonist may also be used as a food additive or preservative  
CC to increase or decrease storage capabilities, fat content or other  
CC nutritional components. The present is a secreted protein of the  
CC invention.  
XX Sequence 1189 AA;  
QY Query Match 76.7%; Score 911; DB 6; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDLPRGLVVAWALSILWFGFTDTPNMDTRKPRVIFGSRTPAFGYTVQQHDSIGNKWLWVGA 60  
Db 1 MDLPRGLVVAWALSILWFGFTDTPNMDTRKPRVIFGSRTPAFGYTVQQHDSIGNKWLWVGA 60  
QY 61 PLETNGYQKTDGVYKCPVIHGNCTKLNIGRVTLSNVSEKDNMRGLSLATNPKNDSFLA 120  
Db 61 PLETNGYQKTDGVYKCPVIHGNCTKLNIGRVTLSNVSEKDNMRGLSLATNPKNDSFLA 120  
QY 121 CSPLWSHECGSSYTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDCSNSIYPWVE 180  
Db 121 CSPLWSHECGSSYTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDCSNSIYPWVE 180  
QY 181 VQHFLINILKKFYTGPGQIQGVVQYGEDVVEHFLNDYRSKDVVEAASHIEQGGTET 240  
Db 181 VQHFLINILKKFYTGPGQIQGVVQYGEDVVEHFLNDYRSKDVVEAASHIEQGGTET 240  
QY 241 RTAFGEIFARSEAFQKGRKGAKKVMIVITDGSHPDLEKVIQQSERDNVTRYAVAVL 300  
Db 241 RTAFGEIFARSEAFQKGRKGAKKVMIVITDGSHPDLEKVIQQSERDNVTRYAVAVL 300  
QY 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKOIVDALGDRIFSLGNTNKNET 360  
Db 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKOIVDALGDRIFSLGNTNKNET 360  
QY 361 SFGELEMSQTFSSHVVEDGVLLGAVGAYDNAGAVLKEFSAGKVIPLRESYLKEPPEELKN 420  
Db 361 SFGELEMSQTFSSHVVEDGVLLGAVGAYDNAGAVLKEFSAGKVIPLRESYLKEPPEELKN 420  
QY 421 HGAYLGYTVTWSVSSRQGRVYVAGAPRNFHTGKVIPLTMNNRSLTHQAMRQGOIGSYF 480  
Db 421 HGAYLGYTVTWSVSSRQGRVYVAGAPRNFHTGKVIPLTMNNRSLTHQAMRQGOIGSYF 480  
QY 481 GSEITTSVDIDGVTVDVLLVGPAMYFNEGRERGVVYVVELKONRFVYNGTLKDSHSYQNA 540  
Db 481 GSEITTSVDIDGVTVDVLLVGPAMYFNEGRERGVVYVVELKONRFVYNGTLKDSHSYQNA 540  
QY 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHNAGATYIFHGFGRGSILKTPKORITASELATG 600

Db 541 RFGSSIASVRLNODSYNDVVVCGAPLEDNHAGAIYIFHGFRGSILKTPQRITASELATG 600  
Qy 601 LOYFGCSIHGQDLNEDGLDLAVGALGNVILWSRPVQINASLHFPESKINIFHRDCK 660  
Db 601 LOYFGCSIHGQDLNEDGLDLAVGALGNVILWSRPVQINASLHFPESKINIFHRDCK 660  
Qy 661 RSGRDATCLAAFLCFTPIFLAPEFQTTTGGIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720  
Db 661 RSGRDATCLAAFLCFTPIFLAPEFQTTTGGIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720  
Qy 721 LISSGOELCBRIHFVLDYADYKPTFSVEYSLEDDHCPMLDDGWPITLAUSYFVWNG 780  
Db 721 LISSGOELCBRIHFVLDYADYKPTFSVEYSLEDDHCPMLDDGWPITLAUSYFVWNG 780  
Qy 781 CNEDERCVPDLVLDARSDLPTAMEYCORVLRKPAQDCSAVTLISFDITTVIIESTRORVAV 840  
Db 781 CNEDERCVPDLVLDARSDLPTAMEYCORVLRKPAQDCSAVTLISFDITTVIIESTRORVAV 840  
Qy 841 EATLERNGENAYSTVLNISOSANLOFASLIQKEDSDGSIETCVNEERLQKQVNCVSYPPF 900  
Db 841 EATLERNGENAYSTVLNISOSANLOFASLIQKEDSDGSIETCVNEERLQKQVNCVSYPPF 900  
Qy 901 RAKAKVAFLRD 911  
Db 901 RAKAKVAFLRD 911

RESULT 11  
ADE63570  
ID ADE63570 standard; protein; 1189 AA.  
XX AC ADE63570;  
XX AC ADE63570;  
XX DT 29-JAN-2004 (first entry)  
XX DE Human Protein Q9URX5, SEQ ID NO 9514.  
XX KW Human; pain; neuronal tissue; Gene therapy;  
XX KW spinal segmental nerve injury; Chronic constriction injury; CCI;  
XX KW spared nerve injury; SNI; Chung.  
XX OS Homo sapiens.  
XX PN WO2003015475-A2.  
XX PD 27-FEB-2003.  
XX PF 14-AUG-2002; 2002WO-US025765.  
XX PR 14-AUG-2001; 2001US-0312147P.  
XX PR 01-NOV-2001; 2001US-0346382P.  
XX PR 26-NOV-2001; 2001US-0333347P.  
XX PA (GEHO ) GEN HOSPITAL CORP.  
XX PA (FARB ) BAYER AG.  
XX PI Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
XX PT New composition comprising two or more isolated polypeptides, useful for  
XX PT preparing a medicament for treating pain in an animal.  
XX PS Claim 1; Page: 1017pp; English.  
XX CC The invention discloses a composition comprising two or more isolated rat  
XX CC or human polynucleotides or a polynucleotide which represents a fragment,  
XX CC derivative or allelic variation of the nucleic acid sequence. Also  
XX CC claimed are a vector comprising the novel polynucleotide, a host cell  
XX CC comprising the vector, a method for identifying a nucleotide sequence  
XX CC which is differentially regulated in an animal subjected to pain and a  
XX CC kit to perform the method, an array, a method for identifying an agent

CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1189 AA;  
Query Match 76.7%; Score 911; DB 7; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDLPRLVAVWALSILWPGFTDTFNMNDRKPRVIFGSRTPAFGTYVQOHDISGNKWLWVGA 60  
Db 1 MDLPRLVAVWALSILWPGFTDTFNMNDRKPRVIFGSRTPAFGTYVQOHDISGNKWLWVGA 60  
Qy 61 PLETVGYQKTDGVYKCPVIHGNCTKLNLRGVTLSNVSRKDNMRGLSLATNPXNSFLA 120  
Db 61 PLETVGYQKTDGVYKCPVIHGNCTKLNLRGVTLSNVSRKDNMRGLSLATNPXNSFLA 120  
Qy 121 CSPLMSHECGSSYTTGMCNRVNFPSKTVAPALQRCQTYMDIIVLDGNSNTPVWE 180  
Db 121 CSPLMSHECGSSYTTGMCNRVNFPSKTVAPALQRCQTYMDIIVLDGNSNTPVWE 180  
Qy 181 VOHFLINILKIFYIGPGQIQGVVQYGEDVHVEFLNDYRSVKDVEAASHIEQGGTET 240  
Db 181 VOHFLINILKIFYIGPGQIQGVVQYGEDVHVEFLNDYRSVKDVEAASHIEQGGTET 240  
Qy 241 RTAFGIEPARSEAFQKGRKGAKKVMIVITDGEHSDSPLEKVIQOQSRDNVTRVAVL 300  
Db 241 RTAFGIEPARSEAFQKGRKGAKKVMIVITDGEHSDSPLEKVIQOQSRDNVTRVAVL 300  
Qy 301 GYVNRGINPETELNEIKYIASDDPDHFFNVDTDEAALKDIDALGDRIIFSLGNTKNET 360  
Db 301 GYVNRGINPETELNEIKYIASDDPDHFFNVDTDEAALKDIDALGDRIIFSLGNTKNET 360  
Qy 361 SFGLEMSQTGFSSHVEDGVLLGAVGAYDMNGAVLKETSAGKVIPLRESYLKPEPEELKN 420  
Db 361 SFGLEMSQTGFSSHVEDGVLLGAVGAYDMNGAVLKETSAGKVIPLRESYLKPEPEELKN 420  
Qy 421 HGAYLGTYVTSVSSRQGRVYVAGAPRHNHKGKVLFTMHNNRSITTHQMRGQOIGSYF 480  
Db 421 HGAYLGTYVTSVSSRQGRVYVAGAPRHNHKGKVLFTMHNNRSITTHQMRGQOIGSYF 480  
Qy 481 GSEITSDVDIGDGVYDVLVVGAPMYFNEGRGRGVYVYELRQNRVYNGTLKDSHSYQNA 540  
Db 481 GSEITSDVDIGDGVYDVLVVGAPMYFNEGRGRGVYVYELRQNRVYNGTLKDSHSYQNA 540  
Qy 541 RFGSSIASVRLNODSYNDVVVCGAPLEDNHAGAIYIFHGFRGSILKTPQRITASELATG 600  
Db 541 RFGSSIASVRLNODSYNDVVVCGAPLEDNHAGAIYIFHGFRGSILKTPQRITASELATG 600  
Qy 601 LOYFGCSIHGQDLNEDGLDLAVGALGNVILWSRPVQINASLHFPESKINIFHRDCK 660  
Db 601 LOYFGCSIHGQDLNEDGLDLAVGALGNVILWSRPVQINASLHFPESKINIFHRDCK 660  
Qy 661 RSGRDATCLAAFLCFTPIFLAPEFQTTTGGIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720  
Db 661 RSGRDATCLAAFLCFTPIFLAPEFQTTTGGIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720

Db 661 RSGRDATCLAAFLCTPIFLAPHFOTTTVIGIRYNATMDERRYTTPRAHLDEGGDRFTNRV 720  
Qy 721 LLSSGQELCERINPHVLDADYVVKVTSVEYSLEDDPHGPMDDGMPETTLRVSVPFVWG 780  
Db 721 LLSSGQELCERINPHVLDADYVVKVTSVEYSLEDDPHGPMDDGMPETTLRVSVPFVWG 780  
Qy 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLSFDTTVFIIESTRQVAV 840  
Db 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLSFDTTVFIIESTRQVAV 840  
Qy 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIIECVNEERLQKQCNVSYPPF 900  
Db 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIIECVNEERLQKQCNVSYPPF 900  
Qy 901 RAKAKVAFRLD 911  
Db 901 RAKAKVAFRLD 911

RESULT 12

AD86584  
ID ADE86584 standard; protein; 1189 AA.

XX AC ADE86584;

XX DT 29-JAN-2004 (first entry)

XX DE Novel human secreted protein #7.

XX KW human; secreted protein; cancer; liver disorder; hepatitis;

XX KW neural disorder; Alzheimer's disease.

XX OS Homo sapiens.

XX US US2003129685-A1.

XX PN 10-JUL-2003.

XX PF 18-APR-2001; 2001US-00836353.

XX PR 28-OCT-1998; 98US-0105971P.

XX PR 27-OCT-1999; 99WO-US025031.

XX PR 19-APR-2000; 2000US-0198407P.

XX PA (NIJ/) NI J.

XX PA (YOUN/) YOUNG P E.

XX PA (KERN/) KENNY J J.

XX PA (OLSE/) OLSEN H S.

XX PA (MOOR/) MOORE P A.

XX PA (WEI/) WEI Y.

XX PA (GRE/) GREENE J M.

XX PA (RUBE/) RUBEN S M.

XX PI Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;

XX PI Ruben SM;

XX DR WPI; 2004-020315/02.

XX DR N-PSDB; AD86586.

XX PT New nucleic acid molecule, useful for preparing a medicament for

XX PT preventing, treating or ameliorating a medical condition e.g. cancer,

XX PT liver disorders or neural disorders.

XX PS Claim 11; SEQ ID NO 35; 380pp; English.

XX CC The invention relates to an isolated nucleic acid sequence, or its

XX CC allelic variant, a fragment of the cDNA sequence, or its fragment,

XX CC domain, epitope or species homologue. The nucleic acid is useful for

XX CC preparing a medicament for preventing, treating or ameliorating a medical

XX CC condition e.g., cancer, liver disorders such as hepatitis or neural

XX CC disorders such as Alzheimer's disease. The present sequence represents

XX CC the amino acid sequence of a novel human secreted protein.

SQ Sequence 1189 AA;  
Query Match 76.7%; Score 911; DB 8; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDLPRGLVAVWALSLEWPGFTDTFNMDTRKPRVIPSRTAFPGYTVOQHDISGNKWLAVGA 60  
Db 1 MDLPRGLVAVWALSLEWPGFTDTFNMDTRKPRVIPSRTAFPGYTVOQHDISGNKWLAVGA 60  
Qy 61 PLENGYQKTDYVKCVIHGNCCTKLNLGRVTLNSVSRKDNMRGLSLATNPKNDSFLA 120  
Db 61 PLENGYQKTDYVKCVIHGNCCTKLNLGRVTLNSVSRKDNMRGLSLATNPKNDSFLA 120  
Qy 121 CSPLWSHECGSSYYTTGMC SRVNSNFRFSKTVA PALQRCQTYMDIVIVL DGSNSIYPWVE 180  
Db 121 CSPLWSHECGSSYYTTGMC SRVNSNFRFSKTVA PALQRCQTYMDIVIVL DGSNSIYPWVE 180  
Qy 181 VQHFLINILKKFYIGPGQIQGVVYQYGEDVWHFHLNDYRSVKDVAASHIEORGTTET 240  
Db 181 VQHFLINILKKFYIGPGQIQGVVYQYGEDVWHFHLNDYRSVKDVAASHIEORGTTET 240  
Qy 241 RTAFGI BFA RSEAFQKGRKGAKVMIVITDGE SHDSPDLEKVIQO SERDNVTYAVAVL 300  
Db 241 RTAFGI BFA RSEAFQKGRKGAKVMIVITDGE SHDSPDLEKVIQO SERDNVTYAVAVL 300  
Qy 301 GYNRRGINPETFLNEIKYI ASDDPHFFNVTD EALKOI VDALGDRIFSL EGTNQTET 360  
Db 301 GYNRRGINPETFLNEIKYI ASDDPHFFNVTD EALKOI VDALGDRIFSL EGTNQTET 360  
Qy 361 SFGLEMSQTGFSSHVVDGVJLGAVGAYDNNGAVL KETSAGKVIPLAESYLKEPPEELKN 420  
Db 361 SFGLEMSQTGFSSHVVDGVJLGAVGAYDNNGAVL KETSAGKVIPLAESYLKEPPEELKN 420  
Qy 421 HGAYLGYTTSVSSRQGRVYVAGAPRPHNTGKVI LFTMHNRS LTIHQAMRGOIGSYF 480  
Db 421 HGAYLGYTTSVSSRQGRVYVAGAPRPHNTGKVI LFTMHNRS LTIHQAMRGOIGSYF 480  
Qy 481 GSEITSVDIDG DGVTDVLLV GAMPYFNEGRERGVYVYELQRNFVYNGTLKDSHSYQNA 540  
Db 481 GSEITSVDIDG DGVTDVLLV GAMPYFNEGRERGVYVYELQRNFVYNGTLKDSHSYQNA 540  
Qy 541 RFGSSIASVRDLNODSYNDVVV GAPLEDNHAGAIYI PHGFRGSILKTPKQRTITASELATG 600  
Db 541 RFGSSIASVRDLNODSYNDVVV GAPLEDNHAGAIYI PHGFRGSILKTPKQRTITASELATG 600  
Qy 601 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILSRPVQINASLHPEPSKINIFHRDCK 660  
Db 601 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILSRPVQINASLHPEPSKINIFHRDCK 660  
Qy 661 RSGRDATCLAAFLCTPIFLAPHFOTTTVIGIRYNATMDERRYTTPRAHLDEGGDRFTNRV 720  
Db 661 RSGRDATCLAAFLCTPIFLAPHFOTTTVIGIRYNATMDERRYTTPRAHLDEGGDRFTNRV 720  
Qy 721 LLSSGQELCERINPHVLDADYVVKVTSVEYSLEDDPHGPMDDGMPETTLRVSVPFVWG 780  
Db 721 LLSSGQELCERINPHVLDADYVVKVTSVEYSLEDDPHGPMDDGMPETTLRVSVPFVWG 780  
Qy 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLSFDTTVFIIESTRQVAV 840  
Db 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLSFDTTVFIIESTRQVAV 840  
Qy 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIIECVNEERLQKQCNVSYPPF 900  
Db 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIIECVNEERLQKQCNVSYPPF 900  
Qy 901 RAKAKVAFRLD 911  
Db 901 RAKAKVAFRLD 911

RESULT 13  
RAB25590

ID AAB25590 standard; protein; 1034 AA.

AC AAB25590;

XX 21-NOV-2000 (first entry)

DT Protein encoded by human secreted protein gene #7 clone HOBY69.

XX Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;

KW rheumatic; dermatological; antiproliferative; antiarteriosclerotic;

KW anticancer; vulnary; antiviral; antibacterial; antifungal;

KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;

KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;

KW Crohn's disease; nephritis; hyperproliferative disorder;

KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;

KW melanoma; lymphoma; wound healing; human; chromosome 15 q22.3-23.

XX Homo sapiens.

OS WO200029435-A1.

XX 25-MAY-2000.

XX 27-OCT-1999; 99WO-US025031.

XX 28-OCT-1998; 98US-0105971P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;

PI Greene JW;

XX WPI; 2000-387742/33.

XX Isolated nucleic acid molecules encoding human secreted proteins are used

PT for the prevention, amelioration and treatment of autoimmune,

PT inflammatory, hyperproliferative and cardiovascular disorders, cancer,

XX wounds, and infectious diseases.

XX Claim 1; Page 678-682; 803pp; English.

XX The present invention relates to 12 secreted human proteins and the

CC nucleotide sequences encoding them. The polynucleotide sequences given in

CC AAB0606-A80623 encode the 12 secreted protein sequences given in

CC AAB25576-B25593. The human secreted proteins have various activities

CC dependent on the tissues in which they are expressed. Examples of the

CC activities of the proteins include: immunosuppressant; anti-inflammatory;

CC antiarthritic; antirheumatic; dermatological; antiproliferative;

CC antiarteriosclerotic; anticancer; vulnary; antiviral; antibacterial;

CC and antifungal activity. The proteins, polypeptides, agonists and

CC antagonists may be used to treat prevent and/or diagnose various disease,

CC disorders and conditions examples of which include: immune disorders e.g.

CC Addison's disease, rheumatoid arthritis, dermatitis, and multiple

CC sclerosis; inflammatory disorders e.g. inflammatory bowel disease,

CC Crohn's disease and nephritis; hyperproliferative disorders such as

CC paraproteinemias and myelopurpura; cardiovascular disorders e.g. coronary

CC arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The

CC proteins and polynucleotide sequences may also be used in wound healing

CC and the treatment of infectious diseases. The human secreted protein gene

CC #7 and protein sequences are represented in sequences AAB0612 and

CC AAB25582. Secreted protein gene #7 is located at position chromosome 15

CC q22.3-23. Sequences AAB0652-A80661 represent genes which are related to

CC the secreted protein gene#7

XX SQ Sequence 1034 AA;

Query Match 68.2%; Score 810; DB 3; Length 1034;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDLPRGLVVAWALSLWPGFTDTFNMDTRKPRVPGSRATFFGYTVQOHDISGNKWLWVGA 60

Db 1 MDLPRGLVVAWALSLWPGFTDTFNMDTRKPRVPGSRATFFGYTVQOHDISGNKWLWVGA 60

QY 61 PLETMGYOKTGDVYKCPVIHGNCTKLNLRVTLNSVSRKDNMRGLSLATNPKDNSFLA 120

DB 61 PLETMGYOKTGDVYKCPVIHGNCTKLNLRVTLNSVSRKDNMRGLSLATNPKDNSFLA 120

QY 121 CSPMSHECGSSYYTTGCMRSVNSNFRFSKTVAPALQRCQTYMDIVIVLOGSNSIYPWVE 180

DB 121 CSPMSHECGSSYYTTGCMRSVNSNFRFSKTVAPALQRCQTYMDIVIVLOGSNSIYPWVE 180

QY 181 VOHFLINILKKFYIGPGQIQGVVQGVGDEVAHEPHLNDYRSVKDVEAASHTEQGGTET 240

DB 181 VOHFLINILKKFYIGPGQIQGVVQGVGDEVAHEPHLNDYRSVKDVEAASHTEQGGTET 240

QY 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGHSDSPDLKXVQQSQRDNVTRYAVAVL 300

DB 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGHSDSPDLKXVQQSQRDNVTRYAVAVL 300

QY 301 GYNRERGINPETFLNEIKYIASDDDDKHFFNVTEAALKIDVADGDRIFSLGNTKNET 360

DB 301 GYNRERGINPETFLNEIKYIASDDDDKHFFNVTEAALKIDVADGDRIFSLGNTKNET 360

QY 361 SFGLEMSQTGFSSHVVEDGVLLGAVGYDMNGAVLKETSGAKVILPRLSYLKEPPEELKN 420

DB 361 SFGLEMSQTGFSSHVVEDGVLLGAVGYDMNGAVLKETSGAKVILPRLSYLKEPPEELKN 420

QY 421 HGAYLGYTVTSVSSROGRVYVAGAPRNHHTGKVLFTMNNRSITIHQMRGQOIGSYF 480

DB 421 HGAYLGYTVTSVSSROGRVYVAGAPRNHHTGKVLFTMNNRSITIHQMRGQOIGSYF 480

QY 481 GSEITSVDIDGVDGVTDLVLLGAPMYFNEGRGRGVYVYELRQNRVYVNGTLKDSYSQNA 540

DB 481 GSEITSVDIDGVDGVTDLVLLGAPMYFNEGRGRGVYVYELRQNRVYVNGTLKDSYSQNA 540

QY 541 RFGSSIASVRDLNODSYNDVVVVGAPLEDNHAGAIYIFHGFRGSILKTPKQRTASELATG 600

DB 541 RFGSSIASVRDLNODSYNDVVVVGAPLEDNHAGAIYIFHGFRGSILKTPKQRTASELATG 600

QY 601 LOYFCCSIHQGLDNLNEDGLIDLAVGALGNVILMSRPVVQINASLHFPSPKINIPIHRDCK 660

DB 601 LOYFCCSIHQGLDNLNEDGLIDLAVGALGNVILMSRPVVQINASLHFPSPKINIPIHRDCK 660

QY 661 RSGRDATCLAFCTPTPIFLAPHFQTTTVGIRYNATMDERRYTTPRAHLDGGDRFTNRAV 720

DB 661 RSGRDATCLAFCTPTPIFLAPHFQTTTVGIRYNATMDERRYTTPRAHLDGGDRFTNRAV 720

QY 721 LLSSQBELCERINFVLTADYVKEVTSVEYSLEDPDHGPMLDGMPTTLRVSPVFWNG 780

DB 721 LLSSQBELCERINFVLTADYVKEVTSVEYSLEDPDHGPMLDGMPTTLRVSPVFWNG 780

QY 781 CNEDEHCVPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTILSPDTTVFIIESTRQVAV 840

DB 781 CNEDEHCVPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTILSPDTTVFIIESTRQVAV 840

QY 841 EATLENRGENAYSTVLNITSQANLQFASLIQKEDSDGSIECVNEERRLQKQVCNYSYPPF 900

DB 841 EATLENRGENAYSTVLNITSQANLQFASLIQKEDSDGSIECVNEERRLQKQVCNYSYPPF 900

QY 901 RAKAKVAFRLD 911

DB 901 RAKAKVAFRLD 911

RESULT 14

ADA27062

ID ADA27062 standard; protein; 1034 AA.

XX ADA27062;

XX 20-NOV-2003 (first entry)

DE Human novel secreted protein from cDNA HOBY69 #2.

XX cytostatic; antiinflammatory; immunomodulator; neuroprotective;



hemostatic; gene therapy; cancer; inflammation; immune disorder; neurological disorder; blood clotting disorder; food additive; preservative; human; secreted protein.

Homo sapiens.

US2003055231-A1.

20-MAR-2003.

29-OCT-2001; 2001US-00984130.

28-OCT-1998; 98US-0105971P.

27-OCT-1999; 99WO-US025031.

19-APR-2000; 2000US-0198407P.

30-OCT-2000; 2000US-0243792P.

18-APR-2001; 2001US-00836353.

(NIJ/J) NI J.

(YOUNG) YOUNG P B.

(Kenny) KENNY J J.

(OLSEN) OLSEN H S.

(MOORE) MOORE P A.

(WEI) WEI Y.

(GREENE) GREENE J M.

(RUBEN) RUBEN S M.

(LIUD) LIU D.

(CROCK) CROCKER P R.

Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM; Ruben SM, Liu D, Crocker PR;

WPI: 2003-567103/53.

N-PSDB; ADA27044.

New human secreted nucleic acid molecules and polypeptides, useful for preventing, treating, or ameliorating a medical condition, such as cancer, inflammation, immune disorders, neurological and blood clotting disorders.

Claim 11; Page 302-305; 454pp; English.

The invention relates to an isolated nucleic molecule that is at least 95% identical to 18 human cDNA sequences representing 12 novel genes encoding secreted proteins or a polynucleotide fragment of the cDNA sequence contained in American Type Culture Collection (ATCC) deposit No. defined in the specification, its species homologue, a variant or allelic variant of the polynucleotide having a polynucleotide capable of hybridising under conditions the polynucleotide, where the polynucleotide does not hybridise under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A or T residues. Also included are recombinant vectors, host cells (for producing the polypeptide), the secreted polypeptide (comprising a sequence that is at least 95% identical to a polypeptide fragment, domain, epitope, full-length protein, variant, allelic variant or species homologue), antibodies that specifically bind to the polypeptides, diagnosing, treating, preventing or ameliorating a medical condition by administering the polynucleotide or the polypeptide, the gene corresponding to the cDNA sequence and identifying an activity in a biological assay (by expressing the cDNA sequence in a cell, isolating the supernatant, and detecting an activity in a biological assay and identifying the protein in the supernatant having the activity). The polypeptides, nucleic acids and antibodies are useful for diagnosing a pathological condition or a susceptibility to a pathological condition, for preventing, treating, or ameliorating a medical condition, such as cancer, inflammation and other immune disorders, neurological and blood clotting disorders (many examples are given in the specification). The nucleic acids are also useful for chromosome identification, radiation hybrid mapping or long-range restriction mapping. The polypeptides and antibodies are useful for providing immunological probes for differential identification of the tissues immunohistochemistry assays. The polypeptide, polynucleotide, agonist or antagonist may also be used as a food additive or preservative to increase or decrease storage capabilities, fat content or other

CC nutritional components. The present is a secreted protein of the  
 CC invention.  
 XX  
 XX  
 SQ Sequence 1034 AA;

Query Match 68.2%; Score 310; DB 6; Length 1034;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDLPRLGVVAWALSMPGFTDTTFNMDTRKPRVPGSRTAFPGYTVQGHDSGNKWLVGA 60  
 Db 1 MDLPRLGVVAWALSMPGFTDTTFNMDTRKPRVPGSRTAFPGYTVQGHDSGNKWLVGA 60

Qy 61 PLETNGYQKTGVDYKCPVHGNCCTKLNIGVTVLSNYSERKDNRLGLSLATNPKDNSFLA 120  
 Db 61 PLETNGYQKTGVDYKCPVHGNCCTKLNIGVTVLSNYSERKDNRLGLSLATNPKDNSFLA 120

Qy 121 CSPLWSHECGSSYTTGMCSSRVNSNFRFSTKVPALQRCQTYMDIVIVLDSGNSIYEWVE 180  
 Db 121 CSPLWSHECGSSYTTGMCSSRVNSNFRFSTKVPALQRCQTYMDIVIVLDSGNSIYEWVE 180

Qy 181 VOHFLINILKFFYIGGQIOGVWQGEDVWHFHLNDYRSVKDVVZAAASHIQRGGTET 240  
 Db 181 VOHFLINILKFFYIGGQIOGVWQGEDVWHFHLNDYRSVKDVVZAAASHIQRGGTET 240

Qy 241 RTAFGIEFAESFAPOKGRGKAKVMIVITDGHSDSPDLKVIQSERDNVTRYAVAVL 300  
 Db 241 RTAFGIEFAESFAPOKGRGKAKVMIVITDGHSDSPDLKVIQSERDNVTRYAVAVL 300

Qy 301 GYNRRGINPETFLNEIKYIASDPDDKHFNVTDEAALKDIVDALGRIIFSLEGTNKEI 360  
 Db 301 GYNRRGINPETFLNEIKYIASDPDDKHFNVTDEAALKDIVDALGRIIFSLEGTNKEI 360

Qy 361 SFGLEMSQTCFSSHVEDGVLLGAVGAYDNGAVLAKETSAGKVIPLRESYLKPEPELKN 420  
 Db 361 SFGLEMSQTCFSSHVEDGVLLGAVGAYDNGAVLAKETSAGKVIPLRESYLKPEPELKN 420

Qy 421 HGAYLGYTVTSSVSSRQGRVYVAGAPRNFHTGKVIPLFMNRRSLTHQAMRQGOIGSYF 480  
 Db 421 HGAYLGYTVTSSVSSRQGRVYVAGAPRNFHTGKVIPLFMNRRSLTHQAMRQGOIGSYF 480

Qy 481 GSITSVDIDGDDGVTDLVLCAPMYENEGRGKVVYVYELRQNFVYNGTLKDSHSTQNA 540  
 Db 481 GSITSVDIDGDDGVTDLVLCAPMYENEGRGKVVYVYELRQNFVYNGTLKDSHSTQNA 540

Qy 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNEAGAIYIFHGFSGSILKTPKQRIITASELATG 600  
 Db 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNEAGAIYIFHGFSGSILKTPKQRIITASELATG 600

Qy 601 LQVFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPVQVQINASLHPEPSKINIFPHDOCK 660  
 Db 601 LQVFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPVQVQINASLHPEPSKINIFPHDOCK 660

Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTPRAHLDGEGDFTNRVAV 720  
 Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTPRAHLDGEGDFTNRVAV 720

Qy 721 LLSGQELCERINPHVLDTDADYKVPVTFVSVEYSLEDDHGPMLDDGHPPTILRVSPVPWNG 780  
 Db 721 LLSGQELCERINPHVLDTDADYKVPVTFVSVEYSLEDDHGPMLDDGHPPTILRVSPVPWNG 780

Qy 781 CNEDEHCVQDLVADARSDLPTAMEYCORVLKPAQDCSATYLSFDTTVFIIESTRQRAVAV 840  
 Db 781 CNEDEHCVQDLVADARSDLPTAMEYCORVLKPAQDCSATYLSFDTTVFIIESTRQRAVAV 840

Qy 841 EATLENREGENAYSTVLNISQSANIQFASLIQKEDSDGSIKEDNERSLQKQCNVSYPPFF 900  
 Db 841 EATLENREGENAYSTVLNISQSANIQFASLIQKEDSDGSIKEDNERSLQKQCNVSYPPFF 900

Qy 901 RAKAKVAFRLD 911  
 Db 901 RAKAKVAFRLD 911



[illegible]

XX 10-APR-2003.  
 PD 02-OCT-2002; 2002WO-US031358.  
 XX 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327342P.  
 PR 09-OCT-2001; 2001US-0327517P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 12-APR-2002; 2002US-0371972P.  
 PR 12-APR-2002; 2002US-0371980P.  
 PR 17-APR-2002; 2002US-0373261P.  
 PR 19-APR-2002; 2002US-0373805P.  
 PR 23-APR-2002; 2002US-0374738P.  
 PR 16-MAY-2002; 2002US-0381101P.  
 PR 17-MAY-2002; 2002US-0381635P.  
 PR 29-MAY-2002; 2002US-0383830P.  
 PR 01-OCT-2002; 2002US-00282839.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Alsobrook JP, Anderson DW, Boldog FL, Burgess CB, Catterton E;  
 PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;  
 PI Kekuda R, Leach MD, Li L, Miller CB, Patturajan M, Rieger DK;  
 PI Rothenberg ME, Shimkets RA, Smithson G, Spytek KA, Taupier RJ;  
 PI Vernet CM, Voss EZ, Zernusen BD, Zhong M;  
 XX  
 DR WPI; 2003-381625/36.  
 DR N-PSDB; ACC72077.  
 XX  
 PT NOVX polypeptides and nucleic acids useful for diagnosing, preventing or  
 PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or  
 PT dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX  
 PS Claim 1; Page 107; 487pp; English.  
 XX  
 CC The present invention relates to novel human NOV proteins and their  
 CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV  
 CC proteins are useful in manufacturing a medicament for treating a syndrome  
 CC associated with a human disease. The NOV proteins and coding sequences  
 CC may be used to diagnose, treat or prevent metabolic disorders such as  
 CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative  
 CC disorders such as Alzheimer's disease or Parkinson's disease, immune  
 CC disorders, haematopoietic disorders and various dyslipidaemias  
 XX  
 SQ Sequence 1120 AA;  
 Query Match 63.4%; Score 753; DB 6; Length 1120;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 159 CQYMDIVIVLDGNSIYPVWEVQHFNLINLKFFVIGCQIQGVGVQGVGVHFEHLND 218  
 DB 90 CQYMDIVIVLDGNSIYPVWEVQHFNLINLKFFVIGCQIQGVGVQGVGVHFEHLND 149  
 QY 219 YRSVKDWEAASHIEQGGTETRTAFGIEFARSEAFQGRGAKKVMIVTDGSHDSP 278  
 DB 150 YRSVKDWEAASHIEQGGTETRTAFGIEFARSEAFQGRGAKKVMIVTDGSHDSP 209  
 QY 279 DLEKVIQOSERNVTRYAVAVLGYNNRGINPETFNLINIKYIASDPDDKHPFNVTDEAL 338  
 DB 210 DLEKVIQOSERNVTRYAVAVLGYNNRGINPETFNLINIKYIASDPDDKHPFNVTDEAL 269

QY 339 KDIVDALGDRIFFSLEGTNRKNETSFGLSEMSQTFSSHVVEDGVLLGAVGAYDMNGAVLKET 398  
 DB 270 KDIVDALGDRIFFSLEGTNRKNETSFGLSEMSQTFSSHVVEDGVLLGAVGAYDMNGAVLKET 329  
 QY 399 SAGKVIPLRESYLKEPPEELKNHGAYLGYTVTSVSSRQGRVYVAGAPRPHNHTKGVILFT 458  
 DB 330 SAGKVIPLRESYLKEPPEELKNHGAYLGYTVTSVSSRQGRVYVAGAPRPHNHTKGVILFT 389  
 QY 459 MHNRSITIHQAMEGQQIGSYFGSEITSVDIDGGDVTDLVLVGAPMYFNEGRERGVVYV 518  
 DB 390 MHNRSITIHQAMEGQQIGSYFGSEITSVDIDGGDVTDLVLVGAPMYFNEGRERGVVYV 449  
 QY 519 ELRQNRFYNGTILKDSHYQNARFGSSIASVRDLNODSYNDVYVGALEDNHAGAIYIFH 578  
 DB 450 ELRQNRFYNGTILKDSHYQNARFGSSIASVRDLNODSYNDVYVGALEDNHAGAIYIFH 509  
 QY 579 GFRGSILKTPKQIRITASELATGLQYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPV 638  
 DB 510 GFRGSILKTPKQIRITASELATGLQYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPV 569  
 QY 639 VQINASLHFEPFSKINIFHRDCKSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMD 698  
 DB 570 VQINASLHFEPFSKINIFHRDCKSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMD 629  
 QY 699 ERYTPRAHLDEGGDRFTNRAVLSSQCELCERINFHVLDTADYVKPVTFSVEYSLEDDP 758  
 DB 630 ERYTPRAHLDEGGDRFTNRAVLSSQCELCERINFHVLDTADYVKPVTFSVEYSLEDDP 689  
 QY 759 HGPMLDDGWPFTTLRVSVFPWNGCNEDEHCVPLDVLARSOLPTAMEXCQVRILRPAQDCS 818  
 DB 690 HGPMLDDGWPFTTLRVSVFPWNGCNEDEHCVPLDVLARSOLPTAMEXCQVRILRPAQDCS 749  
 QY 819 AYTLSEFTTTFILBTRQVAVATLENRGENAVSTVLNISQSANLQFASLIQKEDSDGS 878  
 DB 750 AYTLSEFTTTFILBTRQVAVATLENRGENAVSTVLNISQSANLQFASLIQKEDSDGS 809  
 QY 879 IECVNEERRLQKQVCNVSYPPFFRAKAKVAFELD 911  
 DB 810 IECVNEERRLQKQVCNVSYPPFFRAKAKVAFELD 842  
 RESULT 17  
 AAU19663  
 ID AAU19663 standard; protein; 707 AA.  
 XX  
 AC AAU19663;  
 XX  
 DT 04-DEC-2001 (first entry)  
 XX  
 DE Human novel extracellular matrix protein, Seq ID No 313.  
 XX  
 KW Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;  
 KW antianemic; antirheumatic; antisclerotic; cardiac; vascular;  
 KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;  
 KW human immunodeficiency virus; rheumatoid arthritis; HIV infection; anaemia;  
 KW antiaizheimers; immune/autoimmune disease; HIV infection; multiple sclerosis;  
 KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;  
 KW Sezary syndrome; Gaucher's disease; neurological diseases;  
 KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
 KW cardiac arrest; tachycardia; angina; infection; corneal infections;  
 KW wound healing; immunogen; gene therapy; antisense; food additive.  
 XX  
 OS Homo sapiens.  
 FN WO200155368-A1.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US001348.  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217457P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235634P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249219P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465572/50.  
XX N-PSDB; AAS31234.  
XX  
XX Nucleic acid molecules encoding human secreted extracellular matrix  
XX proteins, used in preventing, treating or ameliorating a disorder, e.g.  
XX Alzheimer's and Parkinson's diseases and cancers.

|    |   |  |
|----|---|--|
| PS | Claim 11; SEQ ID NO 313; 577pp; English.                                  |  |
| XX | The invention relates to isolated nucleic acid molecules encoding novel   |  |
| CC | human secreted extracellular matrix proteins (SPs). The polynucleotides   |  |
| CC | and proteins are used to prevent, treat a medical condition in e.g.       |  |
| CC | humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. For  |  |
| CC | example, disorders associated with decreased expression of SPs. The SP    |  |
| CC | polynucleotide or a vector expressing them may be administered to treat   |  |
| CC | diseases by gene therapy. Antisense molecules may be administered to down |  |
| CC | regulate expression of SPs by binding with the cells own genes and        |  |
| CC | preventing their expression. The polynucleotides may also be used as DNA  |  |
| CC | probes in diagnostic assays. The SPs may also be used as antigens to      |  |
| CC | produce antibodies and to identify modulators (agonists and antagonists)  |  |
| CC | of the SPs. The anti-(SP) antibodies and antagonists may also be used to  |  |
| CC | down regulate expression and activity of SP and as diagnostic agents for  |  |
| CC | detecting the presence of SPs in samples. The disorders include for       |  |
| CC | example: immune/autoimmune diseases (e.g. HIV (human immunodeficiency     |  |
| CC | virus) infections, anaemia, rheumatoid arthritis and multiple sclerosis), |  |
| CC | cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of    |  |
| CC | the breast or liver, Sezary syndrome and Gaucher's disease), neurological |  |
| CC | diseases (e.g. Alzheimer's disease, Parkinson's disease), cardio-         |  |
| CC | /cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina), |  |
| CC | infections caused by bacteria, viruses and fungi and ocular disorders     |  |
| CC | (e.g. corneal infections). Other uses include wound healing, maintenance  |  |
| CC | of organs before transplantation, support of cell culture of primary      |  |
|    | Query Match 41.2%; Score 489; DB 4; Length 707;                           |  |
|    | Best Local Similarity 99.7%; Pred. No. 0;                                 |  |
|    | Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;              |  |
| Qy | 1 MDLPRGLVVAWALSMPGFTDFTNMDTRKPRVPGSRTAFFGVTYQQHDSIGNKMLVGA 60            |  |
| Db | 17 MDLPRGLVVAWALSMPGFTDFTNMDTRKPRVPGSRTAFFGVTYQQHDSIGNKMLVGA 76           |  |
| Qy | 61 PLETNQYQKTDGVYKCPVTHGNCITKLNLRVTLNVSERKDNRLGSLATNPKNSFLA 120           |  |
| Db | 77 PLETNQYQKTDGVYKCPVTHGNCITKLNLRVTLNVSERKDNRLGSLATNPKNSFLA 136           |  |
| Qy | 121 CSPLWSHCEGSSYTTTGMCSRNSNFRSKTVAPALQRCQTYMDIVIVLDGNSLYPWVE 180         |  |
| Db | 137 CSPLWSHCEGSSYTTTGMCSRNSNFRSKTVAPALQRCQTYMDIVIVLDGNSLYPWVE 196         |  |
| Qy | 181 VQHFLNLIKFKYIGPQIQGVQVQYGEDVWHEFHNLNDRSVKDVVAAASHIEQGGTET 240         |  |
| Db | 197 VQHFLNLIKFKYIGPQIQGVQVQYGEDVWHEFHNLNDRSVKDVVAAASHIEQGGTET 256         |  |
| Qy | 241 RTAFGEFARSEAFQGGKRAKKMIVITDGSHDSDPLEKVIQSERDNTVRVAVL 300              |  |
| Db | 257 RTAFGEFARSEAFQGGKRAKKMIVITDGSHDSDPLEKVIQSERDNTVRVAVL 316              |  |
| Qy | 301 GYNNRGINPETFLNEIKYIASDDPKHFNVTDEAALKDIDVDAKGDRIFSLGNTKNET 360         |  |
| Db | 317 GYNNRGINPETFLNEIKYIASDDPKHFNVTDEAALKDIDVDAKGDRIFSLGNTKNET 376         |  |
| Qy | 361 SFGLEMSQTGFSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFFPELQN 420        |  |
| Db | 377 SFGLEMSQTGFSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFFPELQN 436        |  |
| Qy | 421 HGAYLGVTYTSVSSRQGRVYVAGAPRFNHTGKVIILFTMHNRSLLTIHQAMRGQQIGSYF 480      |  |
| Db | 437 HGAYLGVTYTSVSSRQGRVYVAGAPRFNHTGKVIILFTMHNRSLLTIHQAMRGQQIGSYF 496      |  |
| Qy | 481 GSITTSVDIDGDGVTVLLVGAIPMYFNEGRERKGVVYVELRQNFVYNGTLKXDSHYQNA 540       |  |
| Db | 497 GSITTSVDIDGDGVTVLLVGAIPMYFNEGRERKGVVYVELRQNFVYNGTLKXDSHYQNA 556       |  |
| Qy | 541 RFGSSIASVRDLNODSNDVWVGAILEDNHAGAIYIFPHGRSILKTPKORITASBLATG 600        |  |
| Db | 557 RFGSSIASVRDLNODSNDVWVGAILEDNHAGAIYIFPHGRSILKTPKORITASBLATG 616        |  |
| Qy | 601 LOYFGCSIHQQLDNLNEDGLDILAVGALGNVILWSPVQINASLHFEPSKINIFHRDCK 660        |  |
| Db | 617 LOYFGCSIHQQLDNLNEDGLDILAVGALGNVILWSPVQINASLHFEPSKINIFHRDCK 676        |  |

|          |  |  |
|----------|--|--|
| Qy       | 661 RSGRDATCAAFLCFTFIPLAPHFQTTTGI 691                                    |  |
| Db       | 677 RSGRDATCAAFLCFTFIPLAPHFQTTTGI 707                                    |  |
|          | RESULT 18  |  |
| ABP47883 |  |  |
| ID       | ABP47883 standard; protein; 707 AA.                                      |  |
| XX       | AC ABP47883;   |  |
| XX       | XX 23-AUG-2002 (first entry)   |  |
| DT       | Human polypeptide SEQ ID NO 313.   |  |
| DE       | Human; neutropic; neuroprotective; cytostatic; dermatological; virucide; |  |
| XX       | immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnary;   |  |
| XX       | anti-parkinsonian; antiacling; antianemic; antiarthritic; cancer;        |  |
| XX       | antipneumatic; hepatotropic; cerebroprotective; antiinflammatory;        |  |
| XX       | antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;       |  |
| XX       | antiparasitic; cardiac; immune disorder; cardiovascular disorder;        |  |
| XX       | neurological disease; infection; nephrotropic; gene therapy; vaccine.    |  |
| OS       | Homo sapiens.  |  |
| XX       | US2002042386-A1.   |  |
| PN       | 11-APR-2002.   |  |
| PD       | 17-JAN-2001; 2001US-00764870.  |  |
| PF       | 31-JAN-2000; 2000US-0179065P.  |  |
| XX       | 04-FEB-2000; 2000US-0180628P.  |  |
| PR       | 28-JUN-2000; 2000US-0214886P.  |  |
| PR       | 07-JUL-2000; 2000US-0216647P.  |  |
| PR       | 07-JUL-2000; 2000US-0216880P.  |  |
| PR       | 11-JUL-2000; 2000US-0217487P.  |  |
| PR       | 11-JUL-2000; 2000US-0217966P.  |  |
| PR       | 14-JUL-2000; 2000US-0218290P.  |  |
| PR       | 26-JUL-2000; 2000US-0220963P.  |  |
| PR       | 26-JUL-2000; 2000US-0220964P.  |  |
| PR       | 14-AUG-2000; 2000US-0224518P.  |  |
| PR       | 14-AUG-2000; 2000US-0224519P.  |  |
| PR       | 14-AUG-2000; 2000US-0225267P.  |  |
| PR       | 14-AUG-2000; 2000US-0225268P.  |  |
| PR       | 14-AUG-2000; 2000US-0225270P.  |  |
| PR       | 14-AUG-2000; 2000US-0225447P.  |  |
| PR       | 14-AUG-2000; 2000US-0225757P.  |  |
| PR       | 14-AUG-2000; 2000US-0225758P.  |  |
| PR       | 22-AUG-2000; 2000US-0226868P.  |  |
| PR       | 30-AUG-2000; 2000US-0228924P.  |  |
| PR       | 01-SEP-2000; 2000US-0229287P.  |  |
| PR       | 01-SEP-2000; 2000US-0229343P.  |  |
| PR       | 01-SEP-2000; 2000US-0229344P.  |  |
| PR       | 01-SEP-2000; 2000US-0229345P.  |  |
| PR       | 05-SEP-2000; 2000US-0229509P.  |  |
| PR       | 08-SEP-2000; 2000US-0229513P.  |  |
| PR       | 21-SEP-2000; 2000US-0231413P.  |  |
| PR       | 21-SEP-2000; 2000US-0234223P.  |  |
| PR       | 21-SEP-2000; 2000US-0234274P.  |  |
| PR       | 25-SEP-2000; 2000US-0234997P.  |  |
| PR       | 27-SEP-2000; 2000US-0235834P.  |  |
| PR       | 29-SEP-2000; 2000US-0236327P.  |  |
| PR       | 29-SEP-2000; 2000US-0236367P.  |  |
| PR       | 29-SEP-2000; 2000US-0236368P.  |  |
| PR       | 29-SEP-2000; 2000US-0236369P.  |  |
| PR       | 02-OCT-2000; 2000US-0236370P.  |  |
| PR       | 02-OCT-2000; 2000US-0236802P.  |  |
| PR       | 02-OCT-2000; 2000US-0237037P.  |  |
| PR       | 02-OCT-2000; 2000US-0237038P.  |  |
| PR       | 02-OCT-2000; 2000US-0237039P.  |  |
| PR       | 02-OCT-2000; 2000US-0237040P.  |  |

PR 13-OCT-2000; 2000US-0239935P.  
 PR 20-OCT-2000; 2000US-0240950P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0244517P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 XX (ROSE/) ROSEN C A.  
 PA (RUBE/) RUBEN S M.  
 PA (BARA/) BARASH S C.  
 XX  
 PI Rosen CA, Ruben SM, Barash SC;  
 XX  
 DR W21; 2002-470713/50.  
 DR N-PSDB; ABQ66558.  
 XX  
 PT New nucleic acid encoding human proteins, useful for diagnosis, treatment  
 PT and prevention of e.g. osteoporosis, also related polypeptides and  
 PT antibodies.  
 XX  
 PS Claim 11; SEQ ID NO 313; 235pp + Sequence Listing; English.  
 XX  
 CC The invention relates to novel genes (ABQ66521-ABQ66785) and proteins  
 CC (ASP47846-ASP48110) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from USPTO at seqdata.uspto.gov/sequence.html?DocID=999909764870  
 XX  
 SQ Sequence 707 AA;

Query Match 41.2%; Score 489; DB 5; Length 707;  
 Best Local Similarity 99.7%; Pred No. 0;  
 Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MDLPRGLVAVWALSMPGFTDTFMMDTRKPRVPGSRTAFPGYTVQCHDISGNKMLVWGA 60  
 DB 17 MDLPRGLVAVWALSMPGFTDTFMMDTRKPRVPGSRTAFPGYTVQCHDISGNKMLVWGA 76  
 QY 61 PLENTGYOKTGDYKCPVIEHGNCTKLNLRVTLSNVYSERKDNRLGLSLATNPKNLSFLA 120  
 DB 77 PLENTGYOKTGDYKCPVIEHGNCTKLNLRVTLSNVYSERKDNRLGLSLATNPKNLSFLA 136  
 QY 121 CSPLWSHCEGSSYTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWYE 180  
 DB 137 CSPLWSHCEGSSYTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWYE 196  
 QY 181 VQFLINILKFFVIGPQIQGVVQYGEDVWHFPHLNDYRSVKDVAASHIEQRGTTT 240  
 DB 197 VQFLINILKFFVIGPQIQGVVQYGEDVWHFPHLNDYRSVKDVAASHIEQRGTTT 256  
 QY 241 RTAFGIEFARSEAFQGRGKAKKVMIVITDGHSDSPDLEKVIQSERDNTVRYAVVL 300  
 DB 257 RTAFGIEFARSEAFQGRGKAKKVMIVITDGHSDSPDLEKVIQSERDNTVRYAVVL 316  
 QY 301 GYNNRGINPETFLEIKYIASDPDDKHFFNVWDEALKDIDVDAIGRIFFSLEGTNNET 360  
 DB 317 GYNNRGINPETFLEIKYIASDPDDKHFFNVWDEALKDIDVDAIGRIFFSLEGTNNET 376

QY 361 SFGLEMSOTGFSSHHVVDGVLGAVGAYDNGAVLKVTSAGKVIPLRESYLKEFPEELKN 420  
 DB 377 SFGLEMSOTGFSSHHVVDGVLGAVGAYDNGAVLKVTSAGKVIPLRESYLKEFPEELKN 436  
 QY 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSILTIHQAMRGQOIGSYF 480  
 DB 437 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSILTIHQAMRGQOIGSYF 496  
 QY 481 GSEITSVDIDGVTDLVLLVAGAPMFNEGRERGVYVYELRQNRVFNNGTLKXDSHYQNA 540  
 DB 497 GSEITSVDIDGVTDLVLLVAGAPMFNEGRERGVYVYELRQNLVFNNGTLKXDSHYQNA 556  
 QY 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGFSGSILKTPKQRTASLALATG 600  
 DB 557 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGFSGSILKTPKQRTASLALATG 616  
 QY 601 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILWSPVVOINASLHFFPSKINI FHRDCK 660  
 DB 617 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILWSPVVOINASLHFFPSKINI FHRDCK 676  
 QY 661 RSGRDATCLAAFLCFTPIPLAPHFQTTTGI 691  
 DB 677 RSGRDATCLAAFLCFTPIPLAPHFQTTTGI 707  
 RESULT 19  
 ADC10845  
 ID ADC10845 standard; protein; 707 AA.  
 AC ADC10845;  
 XX  
 DT 18-DEC-2003 (first entry)  
 DE Human extracellular matrix protein from gene 38.  
 XX  
 KW Extracellular matrix protein; cytostatic; antibacterial; virucide;  
 KW neuroprotective; gynaecological; gastrointestinal-Gen; cardiant;  
 KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;  
 KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;  
 KW neotropic; antiallergic; cancer; bacterial infection; viral infection;  
 KW neural disorder; immune system disorder; blood disorder;  
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
 KW pulmonary disorder; cardiovascular disorder; renal disorder;  
 KW inflammatory disorder; proliferative disorder; Human.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US2003059875-A1.  
 XX  
 XX 27-MAR-2003.  
 XX  
 XX 19-APR-2002; 2002US-00125540.  
 XX  
 XX 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215133P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 11-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.



CC protein, the gene corresponding to the cDNA sequence, and identifying an  
CC activity in a bioassay (comparing expressing the nucleic acid in  
CC a cell isolating the supernatant, expressing the nucleic acid in  
CC assay and identifying the protein in the supernatant having the  
CC activity). The nucleic acids and proteins display the following  
CC activities: Cytostatic, antibacterial, virucide, Neuroprotective,  
CC Neurological, Gastrointestinal-Gen, Cardiant, Cardiovascular-Gen,  
CC Nephrotropic, Antiinflammatory, Muscular-Gen, Respiratory-Gn.

| Query Match           | 41.2%           | Score 489;         | DB 7;              | length 707;                              |                 |
|-----------------------|-----------------|--------------------|--------------------|--|-----------------|
| Best Local Similarity | 99.7%;          | Pred. No. 0;       |                    |  |                 |
| Matches 689;          | Conservative 0; | Mismatches 2;      | Indels 0;          | Gaps 0;                                  |                 |
| QY                    | 1               | MDLPRLVVAWALS      | LPFGFTDFNM         | DRKPRVPGSRTAFPGYTVQOHDISGNKMLVGA 60      |                 |
| DB                    | 17              | MDLPRLVVAWALS      | LPFGFTDFNM         | DRKPRVPGSRTAFPGYTVQOHDISGNKMLVGA 76      |                 |
| QY                    | 61              | PLETNGYQKTGDVYK    | CVPIHGNCTKMLGRVTL  | SNVSRKDNRMGLSLATNPKDNSFLA 120            |                 |
| DB                    | 77              | PLETNGYQKTGDVYK    | CVPIHGNCTKMLGRVTL  | SNVSRKDNRMGLSLATNPKDNSFLA 136            |                 |
| QY                    | 121             | CSPLWSHCEGSSVYTT   | GMCSRVSNSNFRFSKTV  | APALQRCOTYMDIVVLV                        | DGSNSIYPWVE 180 |
| DB                    | 137             | CSPLWSHCEGSSVYTT   | GMCSRVSNSNFRFSKTV  | APALQRCOTYMDIVVLV                        | DGSNSIYPWVE 196 |
| QY                    | 181             | VQHPLNLIKPFYIGPQOI | GVGVQYGEDVVH       | PHLNDYRSVKDVVEAASHIEORGGTET 240          |                 |
| DB                    | 197             | VQHPLNLIKPFYIGPQOI | GVGVQYGEDVVH       | PHLNDYRSVKDVVEAASHIEORGGTET 256          |                 |
| QY                    | 241             | RTAFGIEFARSEAFQKGR | GAKKVMIVITD        | GSHDSPDLKVIQQSERDNVTRYAVAVL 300          |                 |
| DB                    | 257             | RTAFGIEFARSEAFQKGR | GAKKVMIVITD        | GSHDSPDLKVIQQSERDNVTRYAVAVL 316          |                 |
| QY                    | 301             | GYNRRGINPETTFELNEI | KYIASPDDDKGHFNVT   | DEAALKDIDVALGDRIFSL                      | EGTNKNET 360    |
| DB                    | 317             | GYNRRGINPETTFELNEI | KYIASPDDDKGHFNVT   | DEAALKDIDVALGDRIFSL                      | EGTNKNET 376    |
| QY                    | 361             | SFGLNSQTCFSSHVEDGV | LIGAVGAYDNNGAVL    | KETSAKVPIPLRESYLKEPPELKN 420             |                 |
| DB                    | 377             | SFGLNSQTCFSSHVEDGV | LIGAVGAYDNNGAVL    | KETSAKVPIPLRESYLKEPPELKN 436             |                 |
| QY                    | 421             | HGAYLGYTVTSVSSROGR | VYVAGAPFNNTGKVILFT | MNNRSLTIHQAMRGOQIGSYF 480                |                 |
| DB                    | 437             | HGAYLGYTVTSVSSROGR | VYVAGAPFNNTGKVILFT | MNNRSLTIHQAMRGOQIGSYF 496                |                 |
| QY                    | 481             | GSIITSDVDIGDGVTV   | DVLLGAPMPYFNEG     | RERKGVVYELRKNRFVNGTLKDSHSYQNA 540        |                 |
| DB                    | 497             | GSIITSDVDIGDGVTV   | DVLLGAPMPYFNEG     | RERKGVVYELRQMLFVNGTLKDSHSYQNA 556        |                 |
| QY                    | 541             | RFSSIASVRDLNODS    | VNDVYVWVGA         | PLEDNHAGAIYIPHGFRGSILKTPKORITASELATG 600 |                 |
| DB                    | 557             | RFSSIASVRDLNODS    | VNDVYVWVGA         | PLEDNHAGAIYIPHGFRGSILKTPKORITASELATG 616 |                 |
| QY                    | 601             | LQYFGCSIHQGLD      | LNEDGLDLAV         | ALGNVILMSRPVVOQINASLHPEPSKINIFHRDCK 660  |                 |
| DB                    | 617             | LQYFGCSIHQGLD      | LNEDGLDLAV         | ALGNVILMSRPVVOQINASLHPEPSKINIFHRDCK 676  |                 |
| QY                    | 661             | RSGRDATCLAAE       | LCFTPIFLAPH        | QTTTUVGI 691                             |                 |
| DB                    | 677             | RSGRDATCLAAE       | LCFTPIFLAPH        | QTTTUVGI 707                             |                 |

RESULT 20  
AAU76854  
ID AAU76854 standard; protein; 193 AA.  
XX  
XX AAU76854;  
XX  
XX  
XX 21-MAY-2002 (first entry)  
XX  
XX Human integrin alpha subunit Alpha 11 A domain.  
DE  
DE  
XX  
XX  
XX Human; integrin alpha subunit; A domain; Alpha 11; integrin beta subunit;

|    |   |
|----|---|
| KW | A-like domain; inflammatory disorder; skeletal muscle injury; restenosis; |
| KW | ischaemia-reperfusion injury; immune complex; parasitic disease;          |
| KW | antiinflammatory; vasotropic; antiparasitic; vulnery; gene therapy.       |
| XX |   |
| OS | Homo sapiens.   |
| XX |   |
| PN | W0200209737-A1.   |
| XX |   |
| PD | 07-FEB-2002.  |
| XX |   |
| XX | 31-JUL-2001; 2001WO-US023957.   |
| PF |   |
| XX |   |
| PR | 31-JUL-2000; 2000US-0221950P.   |
| XX |   |
| PR | 11-JAN-2001; 2001US-00758493.   |
| PR | 13-MAR-2001; 2001US-00805354.   |
| XX |   |
| XX |   |
| PA | {GEO } GEN HOSPITAL CORP.   |
| XX |   |
| PI | Arnaout AM, Li R, Xiong J;  |
| XX |   |
| XX | WPI; 2002-1986897/24.   |
| XX |   |
| XX |   |
| PT | Novel high affinity integrin polypeptide useful for treating restenosis   |
| PT | and parasitic diseases, comprises all or part of variant integrin alpha   |
| PT | subunit A domain or variant integrin beta subunit A-like domain.          |
| XX |   |
| PS | Example 2; Fig 5; 55pp; English.  |

CC The invention relates to a high affinity integrin polypeptide comprising  
 CC all or part of a variant integrin alpha subunit A domain or a variant  
 CC integrin beta subunit A-like domain. The polypeptide, preferably the  
 CC Cdlb alpha subunit A domain, where I at residue 332 has been replaced by  
 CC Cdlb alpha subunit A domain, where I at residue 332 has been replaced by  
 CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or  
 CC V at residue 315 and A at residue 320 have been replaced by C, is useful  
 CC for determining if a test compound is a candidate compound for binding to  
 CC Cdlb or for treating an inflammatory disorder, by contacting a test  
 CC compound with the polypeptide and determining if the test compound binds  
 CC to the polypeptide. The integrin subunits are useful for reducing  
 CC skeletal muscle injury, for treating disorders caused by ischaemia-  
 CC reperfusion injury, immune complexes, restenosis and parasitic diseases,  
 CC to purify variant integrin polypeptide ligands and as bait proteins in  
 CC two-hybrid or three-hybrid assays. This sequence represents the human  
 CC integrin alpha subunit Alpha 11 A domain  
 XX  
 SO Sequence 193 AA;

|                       |                 |                     |           |             |
|-----------------------|-----------------|---------------------|-----------|-------------|
| Query Match           | 16.2%;          | Score 193;          | DB 5;     | Length 193; |
| Best Local Similarity | 100.0%;         | Pred. No. 2.1e-182; |           |             |
| Matches 193;          | Conservative 0; | Mismatches 0;       | Indels 0; | Gaps 0      |

|           |                                     |             |                 |                             |                      |                |     |
|-----------|-------------------------------------|-------------|-----------------|-----------------------------|----------------------|----------------|-----|
| Qy        | 159                                 | CQTYMDIVIVL | QSGNSIYPWTEVQHF | FLINILKCFYIGPGQIQGVVQYGEDVW | HFEHLND              | 218            |     |
|           |                                     |             |                 |                             |                      |                |     |
| Db        | 1                                   | CQTYMDIVIVL | QSGNSIYPWTEVQHF | FLINILKCFYIGPGQIQGVVQYGEDVW | HFEHLND              | 60             |     |
|           |                                     |             |                 |                             |                      |                |     |
| Qy        | 219                                 | YRSVKDWW    | VEAAASHIEQRGGT  | ETRTAFGIETAF                | SEAFQKGRKGAKKVMIVITD | GESHDSF        | 278 |
|           |                                     |             |                 |                             |                      |                |     |
| Db        | 61                                  | YRSVKDWW    | VEAAASHIEQRGGT  | ETRTAFGIETAF                | SEAFQKGRKGAKKVMIVITD | GESHDSF        | 120 |
|           |                                     |             |                 |                             |                      |                |     |
| Qy        | 279                                 | DEKVIQ      | QSGRNWTRYAV     | VLGVYNRGINP                 | ETFLNEIKYIASD        | PDQKHFPNVTDEAL | 338 |
|           |                                     |             |                 |                             |                      |                |     |
| Db        | 121                                 | DEKVIQ      | QSGRNWTRYAV     | VLGVYNRGINP                 | ETFLNEIKYIASD        | PDQKHFPNVTDEAL | 180 |
|           |                                     |             |                 |                             |                      |                |     |
| Qy        | 339                                 | KDIVD       | ALGDRI          | FS                          | 351                  |                |     |
|           |                                     |             |                 |                             |                      |                |     |
| Db        | 181                                 | KDIVD       | ALGDRI          | FS                          | 193                  |                |     |
|           |                                     |             |                 |                             |                      |                |     |
| RESULT 21 |                                     |             |                 |                             |                      |                |     |
| AAU76863  |                                     |             |                 |                             |                      |                |     |
| ID        | AAU76863 standard; protein; 193 AA. |             |                 |                             |                      |                |     |
| XX        |                                     |             |                 |                             |                      |                |     |
| XX        |                                     |             |                 |                             |                      |                |     |
| AC        |                                     |             |                 |                             |                      |                |     |
| XX        | AAU76863;                           |             |                 |                             |                      |                |     |
| XX        |                                     |             |                 |                             |                      |                |     |



DT 21-MAY-2002 (first entry)  
DE Human integrin alpha subunit Alpha 11 variant A domain.  
XX  
XX Human; integrin alpha subunit; A domain; Alpha 11; integrin beta subunit;  
KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;  
KW ischaemia-reperfusion injury; immune complex; parasitic disease; mutant;  
KW antiinflammatory; vasotropic; antiparasitic; vulnerary; gene therapy;  
KW mutein.  
XX  
OS Homo sapiens.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Misc-difference 191  
FT /note= "Wild-type Ile substituted by any other amino  
FT acid"  
XX  
XX WO200209737-A1.  
XX  
XX 07-FEB-2002.  
XX  
XX 31-JUL-2001; 2001WO-US023957.  
XX  
XX 31-JUL-2000; 2000US-0221950P.  
PR 11-JAN-2001; 2001US-00758493.  
PR 13-MAR-2001; 2001US-00805354.  
XX  
XX (GENO) GEN HOSPITAL CORP.  
XX  
XX Arnaout AM, Li R, Xiong J;  
PI WPI; 2002-188687/24.  
XX  
XX Novel high affinity integrin polypeptide useful for treating restenosis  
PT and parasitic diseases, comprises all or part of variant integrin alpha  
PT subunit A domain or variant integrin beta subunit A-like domain.  
XX  
XX Claim 57; Page: 55pp; English.  
XX  
XX The invention relates to a high affinity integrin polypeptide comprising  
CC all or part of a variant integrin alpha subunit A domain or a variant  
CC integrin beta subunit A-like domain. The polypeptide, preferably the  
CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by  
CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or  
CC V at residue 315 and A at residue 320 have been replaced by C, is useful  
CC for determining if a test compound is a candidate compound for binding to  
CC CD11b or for treating an inflammatory disorder, by contacting a test  
CC compound with the polypeptide and determining if the test compound binds  
CC to the polypeptide. The integrin subunits are useful for reducing  
CC skeletal muscle injury, for treating disorders caused by ischaemia-  
CC reperfusion injury, immune complexes, restenosis and parasitic diseases,  
CC to purify variant integrin polypeptide ligands and as bait proteins in  
CC two-hybrid or three-hybrid assays. This sequence represents a human  
CC integrin alpha subunit Alpha 11 variant A domain. Note: This variant  
CC sequence is not featured in the specification but has been derived from  
CC the wild-type protein shown in AAU76854  
XX  
XX Sequence 193 AA;  
SQ

Query Match 16.0%; Score 190; DB 5; Length 193;  
Best Local Similarity 100.0%; Pred. No. 2e-179;  
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 159 CQTMDIVIVLDGNSITYPWVEVQHFLNLIKRFYIGPGQIQGVVQYGEDVHFLND 218  
DB 1 CQTMDIVIVLDGNSITYPWVEVQHFLNLIKRFYIGPGQIQGVVQYGEDVHFLND 60  
QY 219 YRSVKDVVEASHIEQGGTETRTAFGIEPARSEAFQGGKKGAKKVMIVITDGSHDSP 278  
DB 61 YRSVKDVVEASHIEQGGTETRTAFGIEPARSEAFQGGKKGAKKVMIVITDGSHDSP 120  
QY 279 DLEKVIQOSRDVNTRYAVAVLGYNNRRGINPETFLNEIKYIASDDPKKFFNVTDEAL 338

DB 121 DLEKVIQOSRDVNTRYAVAVLGYNNRRGINPETFLNEIKYIASDDPKKFFNVTDEAL 180  
QY 339 KDIVDALGDR 348  
DB 181 KDIVDALGDR 190  
RESULT 22  
AAB50087  
ID AAB50087 standard; protein; 1188 AA.  
XX  
XX AAB50087;  
AC AAB50087;  
DT 19-MAR-2001 (first entry)  
XX  
DE Murine A259.  
XX  
XX Human; A259; integrin alpha subunit; fibrosis; liver disease; lung;  
KW kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;  
KW rheumatoid arthritis.  
XX  
XX Mus sp.  
XX  
XX Key Location/Qualifiers  
FH Domain 1..1141  
FT /label= Extracellular\_domain  
FT Peptide 1..22  
FT /label= Signal\_peptide  
FT Protein 23..1188  
FT /label= Mature\_protein  
FT Domain 39..74  
FT /label= Integrin\_alphasubunit\_repeat\_domain\_#1  
FT Domain 115..157  
FT /label= Integrin\_alphasubunit\_repeat\_domain\_#2  
FT Domain 164..345  
FT /label= 1\_domain  
FT Domain 367..392  
FT /label= Integrin\_alphasubunit\_repeat\_domain\_#3  
FT Domain 421..455  
FT /label= Integrin\_alphasubunit\_repeat\_domain\_#4  
FT Domain 478..516  
FT /label= Integrin\_alphasubunit\_repeat\_domain\_#5  
FT Domain 540..575  
FT /label= Integrin\_alphasubunit\_repeat\_domain\_#6  
FT Domain 602..640  
FT /label= Integrin\_alphasubunit\_repeat\_domain\_#7  
FT Domain 1142..1164  
FT /label= Transmembrane\_domain  
FT Domain 1165..1188  
FT /label= Cytoplasmic\_domain  
XX  
XX WO200073339-A1.  
XX  
XX 07-DEC-2000.  
XX  
XX 15-MAY-2000; 2000WO-US013262.  
XX  
XX 28-MAY-1999; 99US-00322790.  
XX 27-APR-2000; 2000US-00561263.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Pan Y, Lora JM;  
XX  
XX WPI; 2001-041142/05.  
XX  
XX N-PSDB; AAC91904, AAC91905.  
XX  
XX Nucleic acid encoding alpha-integrin subunits, useful for treatment and  
PT diagnosis of fibrosis, e.g. of the liver.  
XX  
XX Claim 8; Fig 5; 164pp; English.  
XX

CC The present sequence is murine integrin alpha subunit, A259. A259 is  
 CC homologous with the alpha1 and alpha10 integrin subunits and is  
 CC overexpressed in fibrosis. A259 is implicated in regulation of  
 CC proliferation, differentiation and/or function of many different cell  
 CC types. Inhibitors of A259 activity are useful for the treatment of liver  
 CC disease, particularly fibrosis, and also fibrosis in other organs  
 CC (specifically lung and kidney). In addition, A259 can be used for  
 CC treatment and prevention of cancer, osteoporosis, acute myeloid  
 CC leukaemia, HIV infection, and rheumatoid arthritis  
 XX  
 SQ Sequence 1188 AA;

Query Match 10.3%; Score 122; DB 4; Length 1188;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-111; Mismatches 0; Gaps 0;  
 Matches 122; Conservative 0; Indels 0; Gaps 0;  
 QY 81 GNCYKLNIGRVTLSNVSRKDNRLGSLATNPKNSFLACSPFWSHCEGSSYYTGMCS 140  
 DB 81 GNCYKLNIGRVTLSNVSRKDNRLGSLATNPKNSFLACSPFWSHCEGSSYYTGMCS 140  
 QY 141 RVNSNFRSKTVAPALQRCQTYMDIVIVLDGSSNIIYVVEVQHFLINILKKFYIGPGQIQ 200  
 DB 141 RVNSNFRSKTVAPALQRCQTYMDIVIVLDGSSNIIYVVEVQHFLINILKKFYIGPGQIQ 200  
 QY 201 VG 202  
 DB 201 VG 202

RESULT 23  
 AAU10552  
 ID AAU10552 standard; protein; 1188 AA.  
 AC AAU10552;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Murine A259 polypeptide.  
 KW Mouse; A259; integrin alpha subunit; integrin alpha 10; secreted protein;  
 KW liver disease; fibrosis; lung; kidney; bone associated disorder; blood;  
 KW cartilage associated disorder; haematopoietic disorder; bone marrow;  
 KW immune related disease; apoptotic disorder; neuronal tissue disease;  
 KW neurodegenerative disease; gene therapy; cancer; cytostatic; osteopathic;  
 KW nephrotropic; immunomodulator; anti-inflammatory; neuroprotective;  
 KW antiarthritic; antianemic; antiallergic; antiasthmatic; dermatological;  
 KW antidiabetic; anticonvulsant; antiparkinsonian.  
 XX

Mus musculus.  
 XX  
 FH Key  
 FT Domain Location/Qualifiers  
 FT Peptide 1..1141  
 FT Protein 23..1188 /note= "Signal peptide"  
 FT Domain 39..74 /note= "Mature murine A259"  
 FT Domain 125..157 /note= "Integrin alpha repeat domain"  
 FT Domain 164..345 /note= "Integrin alpha repeat domain"  
 FT Domain 367..392 /note= "I domain or Von Willebrand Factor type A domain"  
 FT Domain 421..455 /note= "Integrin alpha repeat domain"  
 FT Domain 478..516 /note= "Integrin alpha repeat domain"  
 FT Domain 540..575 /note= "Integrin alpha repeat domain"  
 FT Domain 602..640 /note= "Integrin alpha repeat domain"  
 FT Domain

FT Domain 1142..1164 /note= "Transmembrane domain"  
 FT Domain 1165..1188 /note= "Cytoplasmic domain"  
 XX  
 EW W0200181414-A2.  
 XX  
 PD 01-NOV-2001.  
 XX  
 XX 27-APR-2001; 2001MO-US013516.  
 XX  
 XX 27-APR-2000; 2000US-00561263.  
 XX  
 XX (MTLL-) MILLENNIUM PHARM INC.  
 XX  
 XX Pan Y, Lora J;  
 XX  
 DR WPI; 2002-041397/05.  
 DR N-PSDS; AAS16874.  
 XX  
 XX New A259 nucleic acids and polypeptides, which comprise integrin alpha  
 subunit, useful for diagnosing, preventing or treating e.g. liver  
 disease, kidney or lung fibrosis, cancers, blood disorders or immune  
 related diseases.

Claim 9; Fig 5; 168pp; English.

The invention relates to human and murine A259 nucleic acid molecules  
 which encode secreted proteins with homology to integrin alpha subunits,  
 specifically to integrin alpha 10. The A259 polypeptide and nucleic acid  
 are useful for treating liver disease or fibrosis, particularly kidney  
 fibrosis or lung fibrosis. The A259 polypeptide and nucleic acid are also  
 useful for diagnosing, preventing or treating cartilage and bone  
 associated disorders (such as bone cancer, achondroplasia, myeloma,  
 fibrous dysplasia, scoliosis, osteoarthritis, osteosarcoma and  
 osteoporosis), bone marrow, blood and haematopoietic disorders (such as  
 acute myeloid leukaemia, haemophilia, anaemia and thalassaemia), immune  
 related diseases (such as HIV, viral infections, cancers, T cell  
 autoimmune disorders e.g. AIDS, and allergic inflammatory disorders e.g.  
 asthma and psoriasis), apoptotic disorders (such as systemic lupus  
 erythematosus and insulin-dependent diabetes mellitus), diseases of the  
 neuronal tissues (such as epilepsy and muscular dystrophy) and  
 neurodegenerative diseases (such as Parkinson's disease and Huntington's  
 disease). This sequence represents the murine A259 polypeptide

Sequence 1188 AA;

Query Match 10.3%; Score 122; DB 5; Length 1188;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-111; Mismatches 0; Gaps 0;  
 Matches 122; Conservative 0; Indels 0; Gaps 0;  
 QY 81 GNCYKLNIGRVTLSNVSRKDNRLGSLATNPKNSFLACSPFWSHCEGSSYYTGMCS 140  
 DB 81 GNCYKLNIGRVTLSNVSRKDNRLGSLATNPKNSFLACSPFWSHCEGSSYYTGMCS 140  
 QY 141 RVNSNFRSKTVAPALQRCQTYMDIVIVLDGSSNIIYVVEVQHFLINILKKFYIGPGQIQ 200  
 DB 141 RVNSNFRSKTVAPALQRCQTYMDIVIVLDGSSNIIYVVEVQHFLINILKKFYIGPGQIQ 200  
 QY 201 VG 202  
 DB 201 VG 202

RESULT 24

ADE08585  
 ID ADE08585 standard; protein; 360 AA.  
 XX  
 AC ADE08585;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Novel protein (useful for identifying genetic disorders) #740.

KW novel gene; novel protein; tissue marker; molecular weight marker;  
 KW chromosome marker; genetic disorder.  
 XX Unidentified.  
 OS  
 PN WO2003054152-A2.  
 XX  
 XX  
 PD 03-JUL-2003.  
 XX  
 PF 10-DEC-2002; 2002WO-US039555.  
 XX  
 PR 10-DEC-2001; 2001US-0339739P.  
 PR 11-DEC-2001; 2001US-0339453P.  
 PR 14-MAR-2002; 2002US-0365091P.  
 PR 14-MAR-2002; 2002US-0365384P.  
 PR 12-APR-2002; 2002US-0372381P.  
 PR 12-APR-2002; 2002US-0372615P.  
 PR 22-APR-2002; 2002US-00128558.  
 PR 24-APR-2002; 2002US-0376045P.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;  
 PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;  
 PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;  
 XX  
 DR WPI; 2003-569235/53.  
 DR N-PSDB; ADE07674.  
 XX  
 XX New polynucleotides, useful for expressing recombinant proteins for  
 PT analysis, characterization or therapeutic use, or as markers for tissues  
 PT in which the corresponding protein is preferentially expressed.  
 XX  
 PS Claim 20; SEQ ID NO 1651; 1177pp; English.  
 XX  
 CC The invention comprises the amino acid and coding sequences of novel  
 CC proteins. The DNA and protein sequences of the invention are useful as:  
 CC markers for tissues in which the corresponding protein is preferentially  
 CC expressed; as molecular weight markers on gels; as chromosome markers or  
 CC tags; to identify chromosomes or to map related gene positions; and to  
 CC compare with endogenous DNA sequences in patients to identify potential  
 CC genetic disorders. The present amino acid sequence represents a protein  
 CC of the invention.  
 XX  
 SQ Sequence 360 AA;  
 Query Match 8.8%; Score 104; DB 7; Length 360;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-94;  
 Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 251 SEAFQGRGKAKKMMIVITDGHSDSPDLKVIQOSRDNDVTRYAVAVLGYNRRGINP 310  
 DB 232 SEAFQGRGKAKKMMIVITDGHSDSPDLKVIQOSRDNDVTRYAVAVLGYNRRGINP 291  
 QY 311 ETLNBIKYIASPPDDKHFPNTDEALKDIVDALGDRIFSLEG 354  
 DB 292 ETLNBIKYIASPPDDKHFPNTDEALKDIVDALGDRIFSLEG 335  
 RESULT 25  
 ABG66673  
 ID ABG66673 standard; protein; 109 AA.  
 XX  
 AC ABG66673;  
 XX  
 DT 30-AUG-2002 (first entry)  
 XX  
 DE Human novel polypeptide #8.  
 XX  
 KW Human; inflammatory condition; shock; sepsis; immune response; cancer;  
 KW wound healing; central nervous system disease; haematopoiesis;  
 KW peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;

KW myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;  
 KW cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;  
 KW bone degenerative disorder; periodontal disease; reperfusion injury;  
 KW lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;  
 KW allergic condition; thrombolysis; thrombosis; coagulation disorder;  
 KW fungal infection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200244340-A2.  
 XX  
 PD 06-JUN-2002.  
 XX  
 PF 30-NOV-2001; 2001WO-US047004.  
 XX  
 PR 30-NOV-2000; 2000US-00728952.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D;  
 PI Yamazaki V, Ujwal ML, Drmanac RT;  
 XX  
 DR WPI; 2002-508509/54.  
 DR N-PSDB; ABK94897.  
 XX  
 CC Novel nucleic acids and polypeptides for diagnosis, treatment of  
 PT inflammatory, autoimmune, nervous system, myeloid or lymphoid cell  
 PT disorders, cancer and promoting wound healing.  
 XX  
 PS Claim 10; Page 572; 672pp; English.  
 XX  
 CC The invention relates to human novel polynucleotides and associated  
 CC polypeptides. The polynucleotides and polypeptides are useful for  
 CC treating inflammatory conditions such as arthritis, nephritis, Crohn's  
 CC disease, ischaemia-reperfusion injury, shock, sepsis, immune responses  
 CC and cancer and for promoting wound healing. The sequences are used to  
 CC induce the proliferation of neural cells and regeneration of nerve and  
 CC brain tissue, and are useful for the treatment of central and peripheral  
 CC nervous system diseases and neuropathies, such as Alzheimer's disease,  
 CC Parkinson's disease, Huntington's disease and amyotrophic lateral  
 CC sclerosis. The sequences are involved in chemotactic or chemokinetic  
 CC activity. Regulation of haematopoiesis, treatment of myeloid or lymphoid  
 CC cell disorders and platelet disorders such as thrombocytopenia,  
 CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue  
 CC growth, tissue repair, healing of burns, incisions, ulcers, treatment of  
 CC osteoporosis, osteoarthritis, bone degenerative disorders and periodontal  
 CC disease. The sequences of the invention are also useful for gut  
 CC protection or regeneration and treatment of lung or liver fibrosis,  
 CC reperfusion injury in various tissues, immune deficiencies and disorders  
 CC including severe combined immunodeficiency (SCID), bacterial or fungal  
 CC infections, autoimmune disorders e.g. multiple sclerosis and myasthenia  
 CC gravis, allergic conditions such as asthma, thrombolysis or thrombosis  
 CC and coagulation disorders. Sequences ABG66666-ABG66758 represent human  
 CC novel polypeptides of the invention  
 XX  
 SQ Sequence 109 AA;  
 Query Match 7.5%; Score 89; DB 5; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 2e-79;  
 Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDLPRLGLVAVWALSFWPGFTDTFNNMDTRKPRVIPSRTAFPGYTVQOHDISGNKLVVCA 60  
 DB 1 MDLPRLGLVAVWALSFWPGFTDTFNNMDTRKPRVIPSRTAFPGYTVQOHDISGNKLVVCA 60  
 QY 61 PLETTNGYQKTGDVYKCPVHGNCTKLNLG 89  
 DB 61 PLETTNGYQKTGDVYKCPVHGNCTKLNLG 89  
 RESULT 26  
 AAU19822  
 ID AAU19822 standard; protein; 103 AA.

XX AAU19822; 06-DEC-2001 (first entry) Human novel extracellular matrix protein, Seq ID No 472.  
XX AC Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;  
XX DT antianemic; antirheumatic; antisclerotic; cardiant; vascular;  
XX DE cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;  
XX DE antialzheimers; immune/aucoimmune disease; HIV infection; anaemia;  
XX DE human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;  
XX DE cancers; hyperproliferative disorder; breast neoplasm; melanoma;  
XX DE Sezary syndrome; Gaucher's disease; neurological diseases;  
XX DE Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
XX DE cardiac arrest; tachycardia; angina; infection; corneal infections;  
XX DE wound healing; immunogen; gene therapy; antisense; food additive.  
XX OS Homo sapiens.  
XX XX WO200155368-A1.  
XX PD 02-AUG-2001.  
XX XX 17-JAN-2001; 2001WO-US001348.  
XX XX 31-JAN-2000; 2000US-0179065P.  
XX PR 24-FEB-2000; 2000US-0180628P.  
XX PR 02-MAR-2000; 2000US-0186350P.  
XX PR 16-MAR-2000; 2000US-0189874P.  
XX PR 17-MAR-2000; 2000US-0190076P.  
XX PR 18-APR-2000; 2000US-0198123P.  
XX PR 19-MAY-2000; 2000US-0205155P.  
XX PR 01-JUN-2000; 2000US-0209467P.  
XX PR 28-JUN-2000; 2000US-0214896P.  
XX PR 30-JUN-2000; 2000US-0215135P.  
XX PR 07-JUL-2000; 2000US-0216647P.  
XX PR 07-JUL-2000; 2000US-0216880P.  
XX PR 11-JUL-2000; 2000US-0217487P.  
XX PR 11-JUL-2000; 2000US-0217496P.  
XX PR 14-JUL-2000; 2000US-0218290P.  
XX PR 26-JUL-2000; 2000US-0220963P.  
XX PR 26-JUL-2000; 2000US-0220964P.  
XX PR 14-AUG-2000; 2000US-0224518P.  
XX PR 14-AUG-2000; 2000US-0224519P.  
XX PR 14-AUG-2000; 2000US-0225213P.  
XX PR 14-AUG-2000; 2000US-0225214P.  
XX PR 14-AUG-2000; 2000US-0225267P.  
XX PR 14-AUG-2000; 2000US-0225268P.  
XX PR 14-AUG-2000; 2000US-0225270P.  
XX PR 14-AUG-2000; 2000US-0225447P.  
XX PR 14-AUG-2000; 2000US-0225757P.  
XX PR 14-AUG-2000; 2000US-0225758P.  
XX PR 14-AUG-2000; 2000US-0225759P.  
XX PR 18-AUG-2000; 2000US-0226279P.  
XX PR 22-AUG-2000; 2000US-0226681P.  
XX PR 22-AUG-2000; 2000US-0226686P.  
XX PR 22-AUG-2000; 2000US-0227182P.  
XX PR 23-AUG-2000; 2000US-0227009P.  
XX PR 30-AUG-2000; 2000US-0228924P.  
XX PR 01-SEP-2000; 2000US-0229287P.  
XX PR 01-SEP-2000; 2000US-0229343P.  
XX PR 01-SEP-2000; 2000US-0229344P.  
XX PR 01-SEP-2000; 2000US-0229345P.  
XX PR 05-SEP-2000; 2000US-0229509P.  
XX PR 05-SEP-2000; 2000US-0229513P.  
XX PR 06-SEP-2000; 2000US-0230437P.  
XX PR 06-SEP-2000; 2000US-0230438P.  
XX PR 08-SEP-2000; 2000US-0231242P.  
XX PR 08-SEP-2000; 2000US-0231243P.  
XX PR 08-SEP-2000; 2000US-0231244P.  
XX PR 08-SEP-2000; 2000US-0231413P.  
XX PR 08-SEP-2000; 2000US-0231414P.  
XX PR 08-SEP-2000; 2000US-0232080P.  
XX PR 08-SEP-2000; 2000US-0232081P.  
XX PR 12-SEP-2000; 2000US-0233968P.  
XX PR 14-SEP-2000; 2000US-0233977P.  
XX PR 14-SEP-2000; 2000US-0233989P.  
XX PR 14-SEP-2000; 2000US-0233999P.  
XX PR 14-SEP-2000; 2000US-0234000P.  
XX PR 14-SEP-2000; 2000US-023401P.  
XX PR 14-SEP-2000; 2000US-0233063P.  
XX PR 14-SEP-2000; 2000US-0233064P.  
XX PR 14-SEP-2000; 2000US-0233065P.  
XX PR 21-SEP-2000; 2000US-0234223P.  
XX PR 21-SEP-2000; 2000US-0234274P.  
XX PR 25-SEP-2000; 2000US-0234997P.  
XX PR 25-SEP-2000; 2000US-0234998P.  
XX PR 26-SEP-2000; 2000US-0235484P.  
XX PR 27-SEP-2000; 2000US-0235834P.  
XX PR 27-SEP-2000; 2000US-0235836P.  
XX PR 29-SEP-2000; 2000US-0236327P.  
XX PR 29-SEP-2000; 2000US-0236377P.  
XX PR 29-SEP-2000; 2000US-0236368P.  
XX PR 29-SEP-2000; 2000US-0236369P.  
XX PR 29-SEP-2000; 2000US-0236370P.  
XX PR 02-OCT-2000; 2000US-0236802P.  
XX PR 02-OCT-2000; 2000US-0237037P.  
XX PR 02-OCT-2000; 2000US-0237038P.  
XX PR 02-OCT-2000; 2000US-0237039P.  
XX PR 02-OCT-2000; 2000US-0241787P.  
XX PR 20-OCT-2000; 2000US-0241808P.  
XX PR 20-OCT-2000; 2000US-0241809P.  
XX PR 20-OCT-2000; 2000US-0241826P.  
XX PR 01-NOV-2000; 2000US-0244617P.  
XX PR 08-NOV-2000; 2000US-0246474P.  
XX PR 08-NOV-2000; 2000US-0246475P.  
XX PR 08-NOV-2000; 2000US-0246476P.  
XX PR 08-NOV-2000; 2000US-0246477P.  
XX PR 08-NOV-2000; 2000US-0246478P.  
XX PR 08-NOV-2000; 2000US-0246521P.  
XX PR 08-NOV-2000; 2000US-0246524P.  
XX PR 08-NOV-2000; 2000US-0246525P.  
XX PR 08-NOV-2000; 2000US-0246526P.  
XX PR 08-NOV-2000; 2000US-0246527P.  
XX PR 08-NOV-2000; 2000US-0246528P.  
XX PR 08-NOV-2000; 2000US-0246532P.  
XX PR 08-NOV-2000; 2000US-0246609P.  
XX PR 08-NOV-2000; 2000US-0246610P.  
XX PR 08-NOV-2000; 2000US-0246611P.  
XX PR 08-NOV-2000; 2000US-0246613P.  
XX PR 17-NOV-2000; 2000US-0249207P.  
XX PR 17-NOV-2000; 2000US-0249208P.  
XX PR 17-NOV-2000; 2000US-0249209P.  
XX PR 17-NOV-2000; 2000US-0249210P.  
XX PR 17-NOV-2000; 2000US-0249211P.  
XX PR 17-NOV-2000; 2000US-0249212P.  
XX PR 17-NOV-2000; 2000US-0249213P.  
XX PR 17-NOV-2000; 2000US-0249214P.  
XX PR 17-NOV-2000; 2000US-0249215P.  
XX PR 17-NOV-2000; 2000US-0249216P.  
XX PR 17-NOV-2000; 2000US-0249217P.  
XX PR 17-NOV-2000; 2000US-0249218P.  
XX PR 17-NOV-2000; 2000US-0249244P.  
XX PR 17-NOV-2000; 2000US-0249245P.  
XX PR 17-NOV-2000; 2000US-0249264P.  
XX PR 17-NOV-2000; 2000US-0249265P.

PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251398P.  
 PR 06-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WFI; 2001-465572/50.  
 DR N-PSDB; AAS31393.  
 XX  
 PT Nucleic acid molecules encoding human secreted extracellular matrix  
 PT proteins, used in preventing, treating or ameliorating a disorder, e.g.  
 PT Alzheimer's and Parkinson's diseases and cancers.  
 XX  
 PS Claim 11; SEQ ID NO 472; 577pp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding novel  
 CC human secreted extracellular matrix proteins (SPs). The polynucleotides  
 CC and proteins are used to prevent, treat a medical condition in e.g.  
 CC humans, mice, rabbits, goats, cats, dogs, chickens or sheep. For  
 CC example, disorders associated with decreased expression of SPs. The SP  
 CC polynucleotide or a vector expressing them may be administered to treat  
 CC diseases by gene therapy. Antisense molecules may be administered to down  
 CC regulate expression of SPs by binding with the cells own genes and  
 CC preventing their expression. The polynucleotides may also be used as DNA  
 CC probes in diagnostic assays. The SPs may also be used as antigens to  
 CC produce antibodies and to identify modulators (agonists and antagonists)  
 CC of the SPs. The anti-(SP) antibodies and antagonists may also be used to  
 CC down regulate expression and activity of SP and as diagnostic agents for  
 CC detecting the presence of SPs in samples. The disorders include for  
 CC example: immun/autoimmune diseases (e.g. HIV (human immunodeficiency  
 CC virus) infections, anaemia, rheumatoid arthritis and multiple sclerosis),  
 CC cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of  
 CC the breast or liver, Sezary syndrome and Gaucher's disease), neurological  
 CC diseases (e.g. Alzheimer's disease, Parkinson's disease), cardio-  
 CC /cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina),  
 CC infections caused by bacteria, viruses and fungi and ocular disorders  
 CC (e.g. corneal infections). Other uses include wound healing, maintenance  
 CC of organs before transplantation, support of cell culture of primary

Query Match 7.4%; Score 88; DB 4; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-78;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQYMDIVIVLDCGNSIYFWVQVHFLNLIKFKYIGPGQIQGVVQVCGEDVVFHFLND 218  
 Db 8 CQYMDIVIVLDCGNSIYFWVQVHFLNLIKFKYIGPGQIQGVVQVCGEDVVFHFLND 67

Qy 219 YRSVKDVAASHIEQGGTETRTAFGI 246  
 Db 68 YRSVKDVAASHIEQGGTETRTAFGI 95

RESULT 27  
 AAU87675  
 ID AAU87675 standard; protein; 103 AA.  
 XX  
 AC AAU87675;  
 XX

DT 05-JUN-2002 (first entry)  
 XX  
 DE Novel central nervous system protein #585.  
 XX  
 KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
 KW hyperproliferative disorder; neoplasm; cardiovascular disorder;  
 KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;  
 KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
 KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
 KW adenocarcinoma; reproductive system disorder; testicular feminisation;  
 KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;  
 KW respiratory disorder; renal disorder; kidney failure; blood disorder;  
 KW myocardial infarction; wound healing; cell proliferation; skin aging;  
 KW food additive; food preservative; gene therapy.  
 OS Homo sapiens.  
 XX  
 XX WO200155318-A2.  
 PN  
 XX  
 PD 02-AUG-2001.  
 XX  
 XX  
 PF 17-JAN-2001; 2001WO-US001332.  
 XX  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214888P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226682P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 14-SEP-2000; 2000US-0233066P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239393P.  
 PR 13-OCT-2000; 2000US-0239397P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241122P.  
 PR 20-OCT-2000; 2000US-0241178P.  
 PR 20-OCT-2000; 2000US-0241178P.  
 PR 20-OCT-2000; 2000US-0241178P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249246P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.

PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-581633/65.  
 XX N-PSDB; ABK44005.  
 DR  
 XX  
 PT New isolated nucleic acid encoding a protein for diagnosing, preventing, or  
 PT treating or ameliorating medical conditions and used as food additives or  
 PT preservatives.  
 PT  
 XX  
 PS Claim 9; SEQ ID NO 1193; 837pp; English.  
 XX  
 CC The invention describes an isolated nucleic acid molecule (I) encoding a  
 CC novel central nervous system protein. (I) and polypeptides (III) encoded  
 CC by (I), are used to treat a medical condition and in diagnosis of a  
 CC pathological condition. Disorders which are diagnosed or treated include  
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
 CC angioneurosis, nervous system disorders e.g. Alzheimer's disease and  
 CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses  
 CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders  
 CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,  
 CC adenocarcinomas and irritable bowel syndrome, endocrine disorders e.g. diabetes  
 CC disorders e.g. testicular feminisation, disorders at the cellular level e.g.  
 CC leukaemia, disorders involving neovascularisation e.g. malignancies,  
 CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.  
 CC acute kidney failure and blood related disorders e.g. myocardial  
 CC infarction. The polypeptides can also be used to aid wound healing and  
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to  
 CC maintain organs before transplantation, for supporting cell culture of  
 CC primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities, fat content, lipid, protein,  
 CC Query Match 7.4%; Score 88; DB 4; Length 103;  
 CC Best Local Similarity 100.0%; Pred. No. 1.9e-78;  
 CC Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 CQTYNDIVILVLDGNSIYPVVEVOHFLINILKKYIGPGQIQGVGVQGVGDVWHEFLND 218  
 DB 8 CQTYNDIVILVLDGNSIYPVVEVOHFLINILKKYIGPGQIQGVGVQGVGDVWHEFLND 67  
 QY 219 YRSVKDVVEAASHIEQRGGTETRTAFGI 246  
 DB 68 YRSVKDVVEAASHIEQRGGTETRTAFGI 95  
 RESULT 28  
 ID ABP48042  
 ID ABP48042 standard; protein; 103 AA.  
 AC ABP48042;  
 XX  
 XX 23-AUG-2002 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 472.

XX Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide;  
 XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
 XX antiparkinsonian; antiskinking; antianaemic; antiarthritis; cancer;  
 XX antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
 XX antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
 XX antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
 XX neurological disease; infection; neurotropic; gene therapy; vaccine.  
 OS Homo sapiens.  
 XX US2002042386-A1.  
 XX 11-APR-2002.  
 XX 17-JAN-2001; 2001US-00764870.  
 XX 31-JAN-2000; 2000US-0179065P.  
 XX 04-FEB-2000; 2000US-0180628P.  
 XX 28-JUN-2000; 2000US-0214886P.  
 XX 07-JUL-2000; 2000US-0215647P.  
 XX 07-JUL-2000; 2000US-0215680P.  
 XX 11-JUL-2000; 2000US-0217487P.  
 XX 11-JUL-2000; 2000US-0217496P.  
 XX 14-JUL-2000; 2000US-0218290P.  
 XX 26-JUL-2000; 2000US-0220963P.  
 XX 14-AUG-2000; 2000US-0224518P.  
 XX 14-AUG-2000; 2000US-0224519P.  
 XX 14-AUG-2000; 2000US-0225267P.  
 XX 14-AUG-2000; 2000US-0225268P.  
 XX 14-AUG-2000; 2000US-0225270P.  
 XX 14-AUG-2000; 2000US-0225447P.  
 XX 14-AUG-2000; 2000US-0225757P.  
 XX 14-AUG-2000; 2000US-0225758P.  
 XX 22-AUG-2000; 2000US-0226868P.  
 XX 30-AUG-2000; 2000US-0228924P.  
 XX 01-SEP-2000; 2000US-0228287P.  
 XX 01-SEP-2000; 2000US-0229343P.  
 XX 01-SEP-2000; 2000US-0229344P.  
 XX 01-SEP-2000; 2000US-0229345P.  
 XX 05-SEP-2000; 2000US-0229509P.  
 XX 05-SEP-2000; 2000US-0229513P.  
 XX 08-SEP-2000; 2000US-0231413P.  
 XX 21-SEP-2000; 2000US-0234223P.  
 XX 21-SEP-2000; 2000US-0234274P.  
 XX 25-SEP-2000; 2000US-0234997P.  
 XX 29-SEP-2000; 2000US-0235834P.  
 XX 29-SEP-2000; 2000US-0236327P.  
 XX 29-SEP-2000; 2000US-0236367P.  
 XX 29-SEP-2000; 2000US-0236368P.  
 XX 29-SEP-2000; 2000US-0236369P.  
 XX 29-SEP-2000; 2000US-0236370P.  
 XX 02-OCT-2000; 2000US-0236802P.  
 XX 02-OCT-2000; 2000US-0237037P.  
 XX 02-OCT-2000; 2000US-0237038P.  
 XX 02-OCT-2000; 2000US-0237039P.  
 XX 02-OCT-2000; 2000US-0237040P.  
 XX 13-OCT-2000; 2000US-0239935P.  
 XX 20-OCT-2000; 2000US-0240960P.  
 XX 20-OCT-2000; 2000US-0241785P.  
 XX 20-OCT-2000; 2000US-0241809P.  
 XX 01-NOV-2000; 2000US-0244617P.  
 XX 17-NOV-2000; 2000US-0249299P.  
 XX 08-DEC-2000; 2000US-0251856P.  
 XX 08-DEC-2000; 2000US-0251868P.  
 XX 08-DEC-2000; 2000US-0251869P.  
 XX (ROSE/) ROSEN C A.  
 XX (RUBE/) RUBEN S M.  
 XX (BARA/) BARASH S C.  
 XX Rosen CA, Ruben SM, Barash SC;

XX WPI; 2002-470713/50.  
 DR N-ESDS; ABQ66717.  
 XX New nucleic acid encoding human proteins, useful for diagnosis, treatment  
 PT and prevention of e.g. osteoporosis, also related polypeptides and  
 PT antibodies.  
 XX Claim 11; SEQ ID NO 472; 235pp + Sequence Listing; English.  
 XX The invention relates to novel genes (ABQ66521-ABQ66785) and proteins  
 CC (ABP47846-ABP48110) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer, and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from USPTO at seqdata.uspto.gov/sequence.html?docID=999909764870  
 XX SQ Sequence 103 AA;  
 Query Match 7.4%; Score 88; DB 5; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-78;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLLNLKFFYIGPGQIQGVVQYGVGVVQGVVHFLND 218  
 Db 8 CQTYMDIVIVLDGNSIYPWVEVQHFLLNLKFFYIGPGQIQGVVQYGVGVVQGVVHFLND 67  
 Qy 219 YRSVKDVEEAASHIEQSGTETRTAFGI 246  
 Db 68 YRSVKDVEEAASHIEQSGTETRTAFGI 95  
 RESULT 29  
 ADC11004  
 ID ADC11004 standard; protein; 103 AA.  
 AC ADC11004;  
 XX 18-DEC-2003 (first entry)  
 XX Human protein from extracellular matrix gene 38 #2.  
 XX Extracellular matrix protein; cytostatic; antibacterial; virucide;  
 XX neuroprotective; gynaecological; gastrointestinal-Gen; cardiant;  
 XX cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;  
 XX respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;  
 XX nootropic; antiallergic; cancer; bacterial infection; viral infection;  
 XX neural disorder; immune system disorder; blood disorder;  
 XX muscular disorder; reproductive disorder; gastrointestinal disorder;  
 XX pulmonary disorder; cardiovascular disorder; renal disorder;  
 XX inflammatory disorder; proliferative disorder; Human.  
 OS Homo sapiens.  
 XX US2003059875-A1.  
 XX 27-MAR-2003.  
 XX 19-APR-2002; 2002US-00125540.  
 XX 31-JAN-2000; 2000US-0179065P.  
 XX 04-FEB-2000; 2000US-0180628P.





XX Claim 11; SEQ ID NO 472; 235pp; English.

PS The invention relates to an isolated nucleic acid molecule (cDNA)

XX encoding a human extracellular matrix protein, representing one of 161

CC novel genes. Also included are recombinant vectors, host cells

CC (expressing the protein), the extracellular matrix proteins (including

CC their fragments, epitopes and homologues), an isolated antibody that

CC binds specifically to the protein, diagnosing a pathological condition or

CC susceptibility to a pathological condition (comprising determining the

CC presence or absence of a mutation in the nucleic acid and diagnosing a

CC condition based on the presence or absence of the mutation), diagnosing a

CC pathological condition or susceptibility to a pathological condition

CC (comprising determining the presence or amount of expression of the

CC protein in a biological sample and diagnosing a condition based on the

CC presence or amount of expression of the protein), preventing, treating or

CC ameliorating a medical condition by administering the nucleic acid or

CC protein to a mammalian subject, identifying a binding partner to the

CC protein, the gene corresponding to the cDNA sequence, and identifying an

CC activity in a biological assay (comprising expressing the nucleic acid in

CC a cell, isolating the supernatant, detecting an activity in a biological

CC assay and identifying the protein in the supernatant having the

CC activity). The nucleic acids and proteins display the following

CC activities: Cytostatic, antibacterial, Virucide, Neuroprotective,

CC Gynaecological, Gastrointestinal-Gen, Cardiant, Cardiovascular-Gen,

CC Nephrotropic, Antiinflammatory, Muscular-Gen, Respiratory-Gen,

Query Match 7.4%; Score 88; DB 7; Length 103;

Best Local Similarity 100.0%; Pred. No. 1.9e-78;

Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 CQTYMDIVLVGSGNSIYVWVEVQHFLINILKFFYIGPQIQVGVVQYGEDVHHEFLND 218

DB 8 CQTYMDIVLVGSGNSIYVWVEVQHFLINILKFFYIGPQIQVGVVQYGEDVHHEFLND 67

QY 219 YRSVKDVVAAASHIEORGTEGTARTGI 246

DB 68 YRSVKDVVAAASHIEORGTEGTARTGI 95

RESULT 30

ADB32059

ID ADB32059 standard; peptide; 42 AA.

XX ADB32059;

XX 04-DEC-2003 (first entry)

XX alphaA-integrin alpha subunit alphas1.

XX Integrin; alphaV-beta3 integrin; RGD peptide ligand; modulator; agonist;

XX antagonist; alphas1.

XX Unidentified.

XX WO2003067219-A2.

XX 14-AUG-2003.

XX 07-FEB-2003; 2003WO-US003903.

XX 07-FEB-2002; 2002US-0354773P.

XX (GEHO ) GEN HOSPITAL CORP.

XX Arnaout AM;

XX WPI; 2003-663639/62.

XX Screening potential modulators of alphaVbeta3 integrin useful to identify

XX agonists and antagonists uses computer model of three-dimensional

XX structure including a binding site and data from an alphaVbeta3 integrin-

XX ligand complex.

XX Disclosure; Page 5c; 49pp; English.

PS The invention relates to a method for screening test compounds as

XX potential modulators of alphaV-beta3 integrin using a computer model of

CC the three-dimensional structure of alphaV-beta3 integrin which includes a

CC binding site. The model is based on atomic coordinates of defined alphaV-

CC beta3 integrin amino acids obtained from the structure of a complex of

CC alphaV-beta3 integrin with a known 'RGD peptide' ligand. The method is

CC useful to identify alphaV-beta3 integrin ligands that, because they bind

CC to alphaV-beta3 integrin, may be modulators e.g. agonists or antagonist

CC of alphaV-beta3 integrin activity. It is useful to ascertain whether a

CC specific test compound is a potential modulator and especially to greatly

CC reduce numbers of compounds which must be further tested for their

CC ability to modulate alphaV-beta3 integrin activity. The current sequence

CC represents the alphaA-integrin alpha subunit alphas1.

XX Sequence 42 AA;

QY 335 EAALKDIDVALGDRIFSLSGTNKNETSGLEMSQTGSSHHV 376

DB 1 EAALKDIDVALGDRIFSLSGTNKNETSGLEMSQTGSSHHV 42

RESULT 31

ABB72288

ID ABB72288 standard; protein; 545 AA.

XX ABB72288;

XX 04-APR-2002 (first entry)

XX Murine protein isolated from skin cells SEQ ID NO: 500.

XX Human; rat; mouse; skin cell; skin wound; cancer; growth defect;

XX developmental defect; inflammatory disease; dermatological; vulnerary;

XX immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

XX Mus sp.

XX WO200190357-A1.

XX 29-NOV-2001.

XX 24-MAY-2001; 2001WO-NZ000099.

XX 24-MAY-2000; 2000US-0206650P.

XX 25-JUL-2000; 2000US-0221232P.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;

XX Kumble KD;

XX WPI; 2002-122020/16.

XX New polynucleotides and polypeptides encoded by the polynucleotides

XX isolated from skin cells, useful for treating skin wounds, cancers,

XX growth and developmental defects, inflammatory diseases, or for

XX modulating immune responses.

XX Claim 4; Page 305-306; 466pp; English.

XX The present invention provides the protein and coding sequences of cDNAs

CC isolated from human, murine and rat skin cell libraries. The sequences

CC can be used in the development of therapeutic agents useful in the

CC treatment of skin diseases, including skin wounds, cancer, growth

CC defects, developmental defects and inflammatory diseases. The proteins

CC have important roles in the induction of hair growth, cell proliferation



KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
XX WO200153312-A1.  
XX  
XX  
XX PD 26-JUL-2001.  
XX  
XX PF 26-DEC-2000; 2000WO-US034263.  
XX  
XX PR 23-DEC-1999; 99US-00471275.  
XX PR 21-JAN-2000; 2000US-00488725.  
XX PR 25-APR-2000; 2000US-00552317.  
XX PR 20-JUN-2000; 2000US-00598042.  
XX PR 19-JUL-2000; 2000US-00620312.  
XX PR 03-AUG-2000; 2000US-00653450.  
XX PR 14-SEP-2000; 2000US-00662194.  
XX PR 19-OCT-2000; 2000US-00693036.  
XX PR 29-NOV-2000; 2000US-00727344.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
XX PI Zhou F, Goodrich R, Drmanac RT;  
XX  
XX WPI; 2001-442253/47.  
XX N-PSDB; AAI59084.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders such  
XX as central nervous system injuries.  
XX  
XX Example 4; SEQ ID NO 3073; 10078pp; English.  
XX  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
XX encoded polypeptides (AAI38642-AAI42213) with nootropic,  
XX immunosuppressant and cytostatic activity. The polynucleotides are useful  
XX in gene therapy. A composition containing a polypeptide or polynucleotide  
XX of the invention may be used to treat diseases of the peripheral nervous  
XX system, such as peripheral nervous injuries, peripheral neuropathy and  
XX localised neuropathies and central nervous system diseases, such as  
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
XX utilisation of the activities such as: immune system suppression,  
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
XX assays for receptor activity, arthritis and inflammation, leukaemias and  
XX C.N.S disorders. Note: The sequence data for this patent did not form  
XX part of the printed specification  
XX  
XX Sequence 757 AA;  
XX  
XX Query Match 3.2%; Score 38; DB 4; Length 757;  
XX Best Local Similarity 100.0%; Pred. No. 4.4e-28;  
XX Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX QY 738 DTADYVKPVTFSVSESLDDPDHGMLDGWPPTLRVSU 775  
XX ||||||||||||||||||||||||||||||||||||||||  
XX Db 76 DTADYVKPVTFSVSESLDDPDHGMLDGWPPTLRVSU 113  
XX  
XX RESULT 35  
XX ABG12950  
XX ID ABG12950 standard; protein; 117 AA.  
XX  
XX AC ABG12950;  
XX  
XX DT 13-FEB-2002 (first entry)  
XX  
XX DE Human secreted protein ITGAI1 peptide encoded by secreted protein #6.

XX Novel human diagnostic protein #12941.  
XX  
XX DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
XX OS Homo sapiens.  
XX  
XX XX WO200175067-A2.  
XX  
XX PD 11-OCT-2001.  
XX  
XX PF 30-MAR-2001; 2001WO-US008631.  
XX  
XX PR 31-MAR-2000; 2000US-00540217.  
XX PR 23-AUG-2000; 2000US-00649167.  
XX (HYSE-) HYSEQ INC.  
XX  
XX PI Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
XX N-PSDB; AAS77137.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity.  
XX  
XX Claim 20; SEQ ID NO 43309; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
XX and in recombinant production of (II). The polynucleotides are also used  
XX in diagnostics as expressed sequence tags for identifying expressed  
XX genes. (I) is useful in gene therapy techniques to restore normal  
XX activity of (II) or to treat disease states involving (II). (II) is  
XX useful for generating antibodies against it, detecting or quantitating a  
XX polypeptide in tissue, as molecular weight markers and as a food  
XX supplement. (II) and its binding partners are useful in medical imaging  
XX of sites expressing (II). (I) and (II) are useful for treating disorders  
XX involving aberrant protein expression or biological activity. The  
XX polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
XX amino acid sequences of the invention. Note: The sequence data for this  
XX patent did not appear in the printed specification, but was obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 117 AA;  
XX  
XX Query Match 2.9%; Score 34; DB 4; Length 117;  
XX Best Local Similarity 100.0%; Pred. No. 7.9e-25;  
XX Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX QY 1130 FBTSKQEDWQVPIWIIIVGSTLGGLLIALLVIAL 1163  
XX ||||||||||||||||||||||||||||||||||||  
XX Db 42 FBTSKQEDWQVPIWIIIVGSTLGGLLIALLVIAL 75  
XX  
XX RESULT 36  
XX AAB25614  
XX ID AAB25614 standard; peptide; 33 AA.  
XX  
XX AC AAB25614;  
XX  
XX XX 21-NOV-2000 (first entry)  
XX  
XX DT Human secreted protein ITGAI1 peptide encoded by secreted protein #6.  
XX  
XX DE

XX Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;  
 KW antirheumatic; dermatological; antiproliferative; antiarteriosclerotic;  
 KW anticancer; vulnary; antiviral; antibacterial; antifungal;  
 KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;  
 KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;  
 KW Crohn's disease; nephritis; hyperproliferative disorder;  
 KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;  
 KW melanoma; lymphoma; wound healing; human.  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO200029435-A1.  
 PN  
 XX  
 XX 25-MAY-2000.  
 PD  
 XX  
 XX 27-OCT-1999; 99WO-US025031.  
 PF  
 XX  
 XX 28-OCT-1998; 98US-0105971P.  
 PR  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX  
 XX Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;  
 PI Greene JM;  
 PI  
 XX WPI; 2000-387742/33.  
 DR  
 XX  
 XX Isolated nucleic acid molecules encoding human secreted proteins are used  
 XX for the prevention, amelioration and treatment of autoimmune,  
 XX inflammatory, hyperproliferative and cardiovascular disorders, cancer,  
 XX wounds, and infectious diseases.  
 XX  
 XX Disclosure; Page 107; 803pp; English.  
 PS  
 XX  
 XX The present invention relates to 12 secreted human proteins and the  
 CC nucleotide sequences encoding them. The polynucleotide sequences given in  
 CC AAB0606-A80623 encode the 12 secreted protein sequences given in  
 CC AAB25576-B25593. The human secreted proteins have various activities  
 CC dependent on the tissues in which they are expressed. Examples of the  
 CC activities of the proteins include: immunosuppressant; anti-inflammatory;  
 CC antiarthritic; antirheumatic; dermatological; antiproliferative;  
 CC antiarteriosclerotic; anticancer; vulnary; antiviral; antibacterial;  
 CC and antifungal activity. The proteins, polypeptides, agonists and  
 CC antagonists may be used to treat prevent and/or diagnose various disease,  
 CC disorders and conditions examples of which include: immune disorders e.g.  
 CC Addison's disease, rheumatoid arthritis, dermatitis, and multiple  
 CC sclerosis; inflammatory disorders e.g. inflammatory bowel disease,  
 CC Crohn's disease and nephritis; hyperproliferative disorders such as  
 CC paraproteinemias and purpura; cardiovascular disorders e.g. coronary  
 CC arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The  
 CC proteins and polynucleotide sequences may also be used in wound healing  
 CC and the treatment of infectious diseases. The human secreted protein gene  
 CC #6 and protein sequences are represented in sequences AAB0611 and  
 CC AAB25581. Sequences AAB0650-A80651 represent genes related to the  
 CC secreted protein gene#6  
 XX  
 XX Sequence 33 AA;  
 SQ  
 Query Match 2.8%; Score 33; DB 3; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-24;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 64 TNGYQKTDGVYKCPVHGNCTKLNIGRVTLNV 96  
 Db 1 TNGYQKTDGVYKCPVHGNCTKLNIGRVTLNV 33  
 RESULT 37  
 ADA27114  
 ID ADA27114 standard; peptide; 33 AA.  
 XX  
 AC ADA27114;  
 XX

DT 20-NOV-2003 (first entry)  
 XX Human novel secreted protein from gene 7 peptide.  
 DE  
 XX  
 XX Cytostatic; antiinflammatory; immunomodulator; neuroprotective;  
 KW hemostatic; gene therapy; cancer; inflammation; immune disorder;  
 KW neurological disorder; blood clotting disorder; food additive;  
 KW preservative; human; secreted protein.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US2003055231-A1.  
 PN  
 XX  
 XX 20-MAR-2003.  
 PD  
 XX  
 XX 29-OCT-2001; 2001US-00984130.  
 PF  
 XX  
 XX 28-OCT-1998; 98US-0105971P.  
 PR  
 XX 27-OCT-1999; 99WO-US025031.  
 PR 19-APR-2000; 2000US-0198407P.  
 PR 30-OCT-2000; 2000US-0243792P.  
 PR 18-APR-2001; 2001US-00836353.  
 XX  
 XX (NIJ/) NI J.  
 PA (YOUN/) YOUNG P E.  
 PA (KENN/) KENNY J J.  
 PA (OLSE/) OLSEN H S.  
 PA (MOOR/) MOORE P A.  
 PA (WEI/) WEI Y.  
 PA (GREE/) GREENE J M.  
 PA (RUBE/) RUBEN S M.  
 PA (LIUD/) LIU D.  
 PA (CROC/) CROCKER P R.  
 XX  
 XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;  
 PI Ruben SM, Liu D, Crocker PR;  
 PI  
 XX WPI; 2003-567103/53.  
 DR  
 XX New human secreted nucleic acid molecules and polypeptides, useful for  
 PT preventing, treating, or ameliorating a medical condition, such as  
 PT cancer, inflammation, immune disorders, neurological and blood clotting  
 PT disorders.  
 XX  
 PS Disclosure; Page 343; 454pp; English.  
 CC  
 CC The invention relates to an isolated nucleic molecule that is at least  
 CC 95% identical to 18 human cDNA sequences representing 12 novel genes  
 CC encoding secreted proteins or a polynucleotide fragment of the cDNA  
 CC sequence contained in American Type Culture Collection (ATCC) deposit No.  
 CC defined in the specification, its species homologue, a variant or allelic  
 CC variant of the polynucleotide having a polynucleotide capable of  
 CC hybridising under conditions the polynucleotide, where the polynucleotide  
 CC does not hybridise under stringent conditions to a nucleic acid molecule  
 CC having a nucleotide sequence of only A or T residues. Also included are  
 CC recombinant vectors, host cells (for producing the polypeptide), the  
 CC secreted polypeptide (comprising a sequence that is at least 95%  
 CC identical to a polypeptide fragment, domain, epitope, full-length  
 CC protein, variant, allelic variant or species homologue), antibodies that  
 CC specifically bind to the polypeptides, diagnosing, treating, preventing  
 CC or ameliorating a medical condition by administering the polynucleotide  
 CC or the polypeptide, the gene corresponding to the cDNA sequence and  
 CC identifying an activity in a biological assay (by expressing the cDNA  
 CC sequence in a cell, isolating the supernatant, and detecting an activity  
 CC in a biological assay and identifying the protein in the supernatant  
 CC having the activity). The polypeptides, nucleic acids and antibodies are  
 CC useful for diagnosing a pathological condition or a susceptibility to a  
 CC pathological condition, for preventing, treating, or ameliorating a  
 CC medical condition, such as cancer, inflammation and other immune  
 CC disorders, neurological and blood clotting disorders (many examples are  
 CC given in the specification). The nucleic acids are also useful for  
 CC chromosome identification, radiation hybrid mapping or long-range  
 CC restriction mapping. The polypeptides and antibodies are useful for

CC providing immunological probes for differential identification of the  
CC tissues immunohistochemistry assays. The polypeptide, polynucleotide,  
CC agonist or antagonist may also be used as a food additive or preservative  
CC to increase or decrease storage capabilities, fat content or other  
CC nutritional components. The present is a peptide motif from a secreted  
CC protein of the invention.  
XX  
SQ Sequence 33 AA;

Query Match 2.8%; Score 33; DB 6; Length 33;  
Best Local Similarity 100.0%; Pred. No. 2.6e-24;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 TNGYQKTGDYKCPVHGNCTKLNGLGRVTLNV 96  
DB 1 TNGYQKTGDYKCPVHGNCTKLNGLGRVTLNV 33

## RESULT 38

AD86651  
ID ADE86651 standard; peptide; 33 AA.

AC ADE86651;

DT 29-JAN-2004 (first entry)

XX Transmembrane domain.

XX human; secreted protein; cancer; liver disorder; hepatitis;  
KW neural disorder; Alzheimer's disease; transmembrane domain.

OS Homo sapiens.

XX US2003129685-A1.

PN 10-JUL-2003.

XX 18-APR-2001; 2001US-00836353.

XX 28-OCT-1998; 98US-0105971P.

PR 27-OCT-1999; 99WC-US025031.

PR 19-APR-2000; 2000US-0198407P.

XX (NIJ/J) NI J.

PA (YOUNG) YOUNG P E.

PA (KENN) KENNY J J.

PA (OLSE) OLSEN H S.

PA (MOOR) MOORE P A.

PA (WEI) WEI Y.

PA (GREEN) GREENE J M.

PA (RUBEN) RUBEN S M.

XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;  
PI Ruben SM;

XX WPI; 2004-020335/02.

XX New nucleic acid molecule, useful for preparing a medicament for  
PT preventing, treating or ameliorating a medical condition e.g. cancer,  
PT liver disorders or neural disorders.  
XX  
PS Disclosure; SEQ ID NO 102; 380pp; English.  
XX  
CC The invention relates to an isolated nucleic acid sequence, or its  
CC allelic variant, a fragment of the cDNA sequence, or its fragment,  
CC domain, epitope or species homologue. The nucleic acid is useful for  
CC preparing a medicament for preventing, treating or ameliorating a medical  
CC condition e.g., cancer, liver disorders such as hepatitis or neural  
CC disorders such as Alzheimer's disease. The present sequence represents  
CC the amino acid sequence of a transmembrane domain.

XX Sequence 33 AA;

Query Match 2.0%; Score 24; DB 4; Length 24;  
Best Local Similarity 100.0%; Pred. No. 1.7e-15;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 2.8%; Score 33; DB 8; Length 33;  
Best Local Similarity 100.0%; Pred. No. 2.6e-24;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 TNGYQKTGDYKCPVHGNCTKLNGLGRVTLNV 96  
DB 1 TNGYQKTGDYKCPVHGNCTKLNGLGRVTLNV 33

## RESULT 39

AD86651  
ID AAB30927 standard; peptide; 24 AA.

XX AAB30927;

DT 02-APR-2001 (first entry)

XX Peptide derived from a human alpha11 integrin chain.

XX Human; integrin; alpha11 subunit; fibroblast; muscle cell; chondrocyte;  
KW osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;  
KW wound healing; trauma; rheumatoid arthritis; osteoarthritis;  
KW osteoporosis; cartilage damage; bone damage; cartilage.

OS Homo sapiens.

XX WO2000075187-A1.

XX 14-DEC-2000.

XX 31-MAY-2000; 2000WO-SE001135.

PR 03-JUN-1999; 99SE-00002056.

XX (ACTI-) ACTIVE BIOTECH AB.

XX Gullberg D;

XX WPI; 2001-071061/08.

XX Integrin subunit alpha 11 or integrin heterodimer comprising subunit  
PT alpha 11 in association with subunit beta, useful for treating muscle  
PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.  
XX  
XX Claim 23; Page 40; 79pp; English.

XX The present sequence is derived from the cytoplasmic domain of the human  
XX integrin subunit, designated alpha11. The alpha11 polynucleotide and  
XX polypeptide are useful as markers of cell target molecules, such as  
XX fibroblasts, muscle cells, chondrocytes, osteoblasts, mesenchymally  
XX derived cells or stem cells. They are also used for determining the  
XX differential-stage of cells during differentiation, development in  
XX pathological conditions, in tissue regeneration, in transplantation or in  
XX therapeutic and physiological repair of tissues. The pathological  
XX conditions involving subunit alpha11 are selected from damage of cells,  
XX muscle dystrophy, fibrosis, wound healing, trauma, rheumatoid arthritis,  
XX osteoarthritis and osteoporosis. The polypeptide is useful for detecting the  
XX formation of cartilage during embryonic development, for detecting  
XX physiological therapeutic repair of cartilage and muscle, for selection  
XX and analysis, or for sorting, isolating or purification of chondrocytes  
XX and muscle cells, for detecting regeneration of cartilage or chondrocytes  
XX during transplantation of cartilage or chondrocytes during  
XX transplantation of cartilage or chondrocytes, respectively, or of muscle  
XX or muscle cells during transplantation of muscle or muscle cells,  
XX respectively, and for studies of differentiation or chondrocytes or  
XX muscle cells

XX Sequence 24 AA;

QY 1165 KLGFRSARRRRRPGDPTPKVLE 1188  
 DB 1 KLGFRSARRRRRPGDPTPKVLE 24

## RESULT 40

AA76112  
 ID AA76112 standard; protein; 158 AA.  
 XX AC  
 XX AC  
 XX AA76112;  
 XX  
 XX 27-MAR-2000 (first entry)  
 XX  
 XX  
 XX Rat integrin homologue, SEQ ID NO:391.  
 XX  
 XX Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;  
 KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;  
 KW secreted; transmembrane; inflammation; cancer; neurological disease;  
 KW angiogenesis; tumour vascularisation; growth disorder;  
 KW developmental disorder; skin wound; hair follicle disorder;  
 KW anti-inflammatory; cytostatic; neuroprotective; vulnery.  
 XX  
 XX Rattus sp.  
 XX  
 XX WO9955865-A1.  
 XX  
 XX 04-NOV-1999.  
 XX  
 XX 29-APR-1999; 99WO-NZ000051.  
 XX  
 XX 29-APR-1998; 98US-00069726.  
 XX  
 XX 09-NOV-1998; 98US-00188930.  
 XX  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 XX  
 XX Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;  
 XX WPI; 2000-072177/06.  
 XX  
 XX N-PSDB; AAC61820.  
 XX  
 XX Novel polynucleotides useful for the treatment of various conditions  
 XX including wounds and cancer.  
 XX  
 XX Claim 4; Page 221; 235pp; English.

The invention relates to novel nucleic acid sequences derived from rat dermal papilla, human keratinocytes and neonatal foreskin fibroblasts, and mouse embryonic skin, keratinocyte stem cells and transit amplifying cells. Polypeptides of the invention may be used to treat inflammation, cancer and neurological diseases. The proteins may be used to stimulate the growth and motility of keratinocytes, to inhibit the growth of cancer cells, to modulate angiogenesis and tumour vascularisation, to modulate skin inflammation, to modulate epithelial cell growth and to inhibit binding of HIV-1 to leukocytes. The invention may also be used to treat growth and developmental defects, skin wounds and hair follicle disorders. Sequences AA75942-Y76123 represent polypeptides encoded by cDNA sequences derived from several mouse, rat or human skin cell types. Sequences AA75942-Y75947, AA76020-Y76021, AA76094-Y76104 and AA76119 are proteins with an N-terminal signal sequence, indicating that they are secreted. Sequences AA75986-Y75989, AA76061-Y76071, AA76106-Y76109 and AA76121-Y76122 are proteins with one or more putative transmembrane domains

Sequence 158 AA;

Query Match 1.8%; Score 21; DB 3; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-12;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 RPVVQINASLHPEPSKINIFH 656  
 DB 57 RPVVQINASLHPEPSKINIFH 77

## RESULT 41

AA56051  
 ID AAB56051 standard; protein; 158 AA.  
 XX AC  
 XX AAB56051;  
 XX  
 XX 08-MAR-2001 (first entry)  
 XX  
 XX Skin cell protein, SEQ ID NO: 391.  
 XX  
 XX Rat; skin cell; cytostatic; antiinflammatory; anti-HIV; nootropic;  
 KW neuroprotective; vulnery; immunomodulatory; vaccine;  
 KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;  
 KW inflammation; neurological disease.  
 XX  
 XX Rattus sp.  
 XX  
 XX WO200069884-A2.  
 XX  
 XX 23-NOV-2000.  
 XX  
 XX 15-MAY-2000; 2000WO-NZ000075.  
 XX  
 XX 14-MAY-1999; 99US-00312283.  
 XX  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 XX  
 XX Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD;  
 XX Murison JG;  
 XX WPI; 2001-007495/01.  
 XX  
 XX N-PSDB; AAC99753.  
 XX  
 XX New isolated polynucleotide used in the identification of genetic  
 XX disorders and encoding polypeptides used for treating inflammatory  
 XX disease, cancer and neurological diseases.  
 XX  
 XX Claim 4; Page 296-297; 352pp; English.

The present sequence is a polypeptide which is expressed in mammalian skin cells. The polypeptide is useful for stimulating keratinocyte growth and motility, inhibiting the growth of cancer cells, modulating angiogenesis, inhibiting angiogenesis and vascularisation of tumours, modulating skin inflammation, stimulating the growth of epithelial cells, inhibiting the binding of human immunodeficiency virus (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and neurological diseases. The polynucleotide can be used as a marker, in the identification of genetic disorders, and for the design of oligonucleotides for examining expression patterns

Query Match 1.8%; Score 21; DB 4; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-12;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 RPVVQINASLHPEPSKINIFH 656  
 DB 57 RPVVQINASLHPEPSKINIFH 77

## RESULT 42

AB72251  
 ID ABB72251 standard; protein; 158 AA.  
 XX AC  
 XX ABB72251;  
 XX  
 XX 04-APR-2002 (first entry)  
 XX  
 XX Rat protein isolated from skin cells SEQ ID NO: 391.  
 XX  
 XX



KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;  
 KW developmental defect; inflammatory disease; dermatological; vulnary;  
 KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

XX Rattus sp.

XX WO200190357-A1.

XX 29-NOV-2001.

XX 24-MAY-2001; 2001WO-NZ000099.

XX 24-MAY-2000; 2000US-0206650P.

XX 25-JUL-2000; 2000US-0221232P.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Watson JD, Strachan L, Sleeman M, Orrust R, Murison JG;  
 PI Kumble KD;

XX WPI; 2002-122020/16.

XX New polynucleotides and polypeptides encoded by the polynucleotides  
 PT isolated from skin cells, useful for treating skin wounds, cancers,  
 PT growth and developmental defects, inflammatory diseases, or for  
 PT modulating immune responses.

XX Example 2; Page 246; 466pp; English.

XX The present invention provides the protein and coding sequences of cDNAs  
 CC isolated from human, murine and rat skin cell libraries. The sequences  
 CC can be used in the development of therapeutic agents useful in the  
 CC treatment of skin diseases, including skin wounds, cancer, growth  
 CC defects, developmental defects and inflammatory diseases. The proteins  
 CC have important roles in the induction of hair growth, cell proliferation  
 CC and cell-cell interaction, in maintaining tissue integrity, in wound  
 CC healing and in modulating immune responses. The present sequence is a  
 CC polypeptide of the invention

XX Sequence 158 AA;

Query Match 1.8%; Score 21; DB 5; Length 158;

Best Local Similarity 100.0%; Pred. No. 8.1e-12; Indels 0; Gaps 0;

Matches 21; Conservative 0; Mismatches 0;

QY 636 RPVQINASLHPEPSKINIEH 656

DB 57 RPVQINASLHPEPSKINIEH 77

RESULT 43

AAAY76111

ID AAAY76111 standard; protein; 85 AA.

XX AC

XX AAAY76111;

XX 27-MAR-2000 (first entry)

XX Rat integrin homologue, SEQ ID NO:390.

XX Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;  
 KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;  
 KW secreted; transmembrane; inflammation; cancer; neurological disease;  
 KW angiogenesis; tumour vascularisation; growth disorder;  
 KW developmental disorder; skin wound; hair follicle disorder;  
 KW anti-inflammatory; cytostatic; neuroprotective; vulnary.

XX Rattus sp.

XX WO9955865-A1.

XX 04-NOV-1999.

XX

PF 29-APR-1999; 99WO-NZ000051.  
 XX 29-APR-1998; 98US-00069726.  
 PR 09-NOV-1998; 98US-00188930.  
 XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Strachan L, Sleeman M, Watson JD, Orrust R, Kumble A, Murison JG;  
 XX WPI; 2000-072177/06.

XX N-PSDB; AA261819.

XX Novel polynucleotides useful for the treatment of various conditions  
 PT including wounds and cancer.

XX Claim 4; Page 221; 235pp; English.

XX The invention relates to novel nucleic acid sequences derived from rat  
 CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,  
 CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying  
 CC cells. Polypeptides of the invention may be used to treat inflammation,  
 CC cancer and neurological diseases. The proteins may be used to stimulate  
 CC the growth and motility of keratinocytes, to inhibit the growth of cancer  
 CC cells, to modulate angiogenesis and tumour vascularisation, to modulate  
 CC skin inflammation, to modulate epithelial cell growth and to inhibit  
 CC binding of HIV-1 to leukocytes. The invention may also be used to treat  
 CC growth and developmental defects, skin wounds and hair follicle  
 CC disorders. Sequences AAAY75942-Y76123 represent polypeptides encoded by  
 CC cDNA sequences derived from several mouse, rat or human skin cell types.  
 CC Sequences AAAY75942-Y75947, AAAY76020-Y76021, AAAY76094-Y76104 and AAAY76119  
 CC are proteins with an N-terminal signal sequence, indicating that they are  
 CC secreted. Sequences AAAY75986-Y75989, AAAY76061-Y76071, AAAY76106-Y76109 and  
 CC AAAY76121-Y76122 are proteins with one or more putative transmembrane  
 CC domains

XX Sequence 85 AA;

Query Match 1.4%; Score 17; DB 3; Length 85;

Best Local Similarity 100.0%; Pred. No. 4.4e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 NGTLKDSHSYQNAFQS 544

DB 28 NGTLKDSHSYQNAFQS 44

RESULT 44

AAAB56050

ID AAAB56050 standard; protein; 85 AA.

XX AC

XX AAAB56050;

XX 08-MAR-2001 (first entry)

XX Skin cell protein, SEQ ID NO: 390.

XX Rat; skin cell; cytostatic; antiinflammatory; anti-HIV; nootropic;  
 KW neuroprotective; vulnary; immunomodulatory; vaccine; inhibition;  
 KW keratinocyte growth stimulation; cancer; angiogenesis; inhibition;  
 KW inflammation; neurological disease.

XX Rattus sp.

XX WO200069884-A2.

XX 23-NOV-2000.

XX 15-MAY-2000; 2000WO-NZ0000075.

XX 14-MAY-1999; 99US-00312283.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX

PI Watson JD, Strachan L, Orrust R, Sleeman M, Kumble KD;  
PI Murison JG;  
DR WPI; 2001-007495/01.  
DR N-PSDB; AAC99752.  
XX  
XX New isolated polynucleotide used in the identification of genetic  
PT disorders and encoding polypeptides used for treating inflammatory  
PT disease, cancer and neurological diseases.  
XX  
XX Claim 4; Page 296; 352pp; English.  
XX  
XX The present sequence is a polypeptide which is expressed in mammalian  
CC skin cells. The polypeptide is useful for stimulating keratinocyte growth  
CC and motility, inhibiting the growth of cancer cells, modulating  
CC angiogenesis, inhibiting angiogenesis and vascularisation of tumours,  
CC modulating skin inflammation, stimulating the growth of epithelial cells,  
CC inhibiting the binding of human immunodeficiency virus (HIV)-1 to  
CC leukocytes, and treating inflammatory disease, cancer and neurological  
CC diseases. The polynucleotide can be used as a marker, in the  
CC identification of genetic disorders, and for the design of  
CC oligonucleotides for examining expression patterns  
XX  
XX  
SQ Sequence 85 AA;  
Query Match 1.4%; Score 17; DB 4; Length 85;  
Best Local Similarity 100.0%; Pred. No. 4.4e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 528 NGTLKDSHSYQNARFGS 544  
Db 28 NGTLKDSHSYQNARFGS 44  
RESULT 45  
ABBY72250  
ID ABB72250 standard; protein; 85 AA.  
XX AC ABB72250;  
XX  
XX 04-APR-2002 (first entry)  
XX  
XX Rat protein isolated from skin cells SEQ ID NO: 390.  
XX  
XX Human; rat; mouse; skin cell; skin wound; cancer; growth defect;  
KW developmental defect; inflammatory disease; dermatological; vulnary;  
KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.  
XX  
XX Rattus sp.  
XX  
XX WO200190357-A1.  
XX  
XX 29-NOV-2001.  
XX  
XX 24-MAY-2001; 2001WO-NZ000099.  
XX  
XX 24-MAY-2000; 2000US-0206650P.  
XX  
XX 25-JUL-2000; 2000US-0221232P.  
XX  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
XX Watson JD, Strachan L, Sleeman M, Orrust R, Murison JG;  
PI Kumble KD;  
XX  
XX WPI; 2002-122020/16.  
XX  
XX New polynucleotides and polypeptides encoded by the polynucleotides  
PT isolated from skin cells, useful for treating skin wounds, cancers,  
PT growth and developmental defects, inflammatory diseases, or for  
PT modulating immune responses.  
XX  
XX Example 2; Page 245; 466pp; English.  
XX

CC The present invention provides the protein and coding sequences of cDNAs  
CC isolated from human, murine and rat skin cell libraries. The sequences  
CC can be used in the development of therapeutic agents useful in the  
CC treatment of skin diseases, including skin wounds, cancer, growth  
CC defects, developmental defects and inflammatory diseases. The proteins  
CC have important roles in the induction of hair growth, cell proliferation  
CC and cell-cell interaction, in maintaining tissue integrity, in wound  
CC healing and in modulating immune responses. The present sequence is a  
XX polypeptide of the invention  
XX  
SQ Sequence 85 AA;  
Query Match 1.4%; Score 17; DB 5; Length 85;  
Best Local Similarity 100.0%; Pred. No. 4.4e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 528 NGTLKDSHSYQNARFGS 544  
Db 28 NGTLKDSHSYQNARFGS 44  
RESULT 46  
AAU19634  
ID AAU19634 standard; protein; 148 AA.  
XX AC AAU19634;  
XX  
XX 04-DEC-2001 (first entry)  
XX  
XX Human novel extracellular matrix protein, Seq ID No 284.  
XX  
XX Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;  
KW antianemic; antirheumatic; antisclerotic; cardiac; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;  
KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;  
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;  
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;  
KW Sezary syndrome; Gaucher's disease; neurological diseases;  
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
KW cardiac arrest; tachycardia; angina; infection; corneal infections;  
KW wound healing; immunogen; gene therapy; antisense; food additive.  
XX  
XX Homo sapiens.  
XX  
XX WO200155368-A1.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001348.  
XX  
XX 31-JAN-2000; 2000US-0179063P.  
XX  
XX 04-FEB-2000; 2000US-0180628P.  
XX  
XX 24-FEB-2000; 2000US-0184664P.  
XX  
XX 02-MAR-2000; 2000US-0186350P.  
XX  
XX 16-MAR-2000; 2000US-0189874P.  
XX  
XX 17-MAR-2000; 2000US-0190078P.  
XX  
XX 18-APR-2000; 2000US-0198123P.  
XX  
XX 19-MAY-2000; 2000US-0205515P.  
XX  
XX 28-JUN-2000; 2000US-0209467P.  
XX  
XX 28-JUN-2000; 2000US-0214886P.  
XX  
XX 30-JUN-2000; 2000US-0215135P.  
XX  
XX 07-JUL-2000; 2000US-0216647P.  
XX  
XX 07-JUL-2000; 2000US-0218880P.  
XX  
XX 11-JUL-2000; 2000US-0217487P.  
XX  
XX 11-JUL-2000; 2000US-0217496P.  
XX  
XX 14-JUL-2000; 2000US-0220963P.  
XX  
XX 26-JUL-2000; 2000US-0220964P.  
XX  
XX 14-AUG-2000; 2000US-0224518P.  
XX  
XX 14-AUG-2000; 2000US-0224519P.  
XX  
XX 14-AUG-2000; 2000US-0225213P.  
XX  
XX 14-AUG-2000; 2000US-0225214P.  
XX  
XX 14-AUG-2000; 2000US-0225266P.



CC /cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina),  
CC infections caused by bacteria, viruses and fungi and ocular disorders  
CC (e.g. corneal infections). Other uses include wound healing, maintenance  
CC of organs before transplantation, support of cell culture of primary

Query Match 1.38; Score 15; DB 4; Length 148;

Best Local Similarity 100.0%; Pred. No. 6.9e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGDSNIPW 178

Db 47 DIVVLGDSNIPW 61

RESULT 47

AAU19794

ID AAU19794 standard; protein; 148 AA.

XX AC AAU19794;

XX DT 04-DEC-2001 (first entry)

DE Human novel extracellular matrix protein, Seq ID No 444.

XX Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;  
XX anti-anemic; antirheumatic; antisclerotic; cardiac; vascular;  
XX cerebroprotective; thrombolytic; antimicrobial; HIV infection; anaemia;  
XX antialzheimers; immune/autoimmune disease; HIV infection; melanoma;  
XX human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;  
XX cancers; hyperproliferative disorder; breast neoplasm; melanoma;  
XX Sezary syndrome; Gaucher's disease; neurological diseases;  
XX Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
XX cardiac arrest; tachycardia; angina; infection; corneal infections;  
XX wound healing; immunogen; gene therapy; antisense; food additive.

XX Homo sapiens.

XX WO200155368-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001348.

XX 31-JAN-2000; 2000US-0179065P.

XX 24-FEB-2000; 2000US-0180628P.

XX 02-MAR-2000; 2000US-0184654P.

XX 16-MAR-2000; 2000US-0186350P.

XX 17-MAR-2000; 2000US-0189874P.

XX 18-APR-2000; 2000US-0190076P.

XX 19-MAY-2000; 2000US-0198233P.

XX 07-JUN-2000; 2000US-0205515P.

XX 28-JUN-2000; 2000US-0209467P.

XX 30-JUN-2000; 2000US-0214886P.

XX 07-JUL-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 11-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 14-JUL-2000; 2000US-0217496P.

XX 26-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 14-AUG-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241825P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0251989P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI: 2001-465572/50.  
XX N-PSDB; AAS31365.  
XX  
XX Nucleic acid molecules encoding human secreted extracellular matrix  
XX proteins, used in preventing, treating or ameliorating a disorder, e.g.  
XX Alzheimer's and Parkinson's diseases and cancers.  
XX  
XX Claim 11; SEQ ID NO 444; 577bp; English.  
XX  
XX The invention relates to isolated nucleic acid molecules encoding novel  
XX human secreted extracellular matrix proteins (SPs). The polynucleotides  
XX and proteins are used to prevent, treat a medical condition in e.g.  
XX humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. For  
XX example, disorders associated with decreased expression of SPs. The SP  
XX polynucleotide or a vector expressing them may be administered to treat  
XX diseases by gene therapy. Antisense molecules may be administered to down  
XX regulate expression of SPs by binding with the cells own genes and  
XX preventing their expression. The polynucleotides may also be used as DNA  
XX probes in diagnostic assays. The SPs may also be used as antigens to  
XX produce antibodies and to identify modulators (agonists and antagonists)  
XX of the SPs. The anti-(SP) antibodies and antagonists may also be used to  
XX detecting the presence of SPs in samples. The disorders include for  
XX example: immune/autoimmune diseases (e.g. HIV (human immunodeficiency  
XX virus) infections, anaemia, rheumatoid arthritis and multiple sclerosis),  
XX cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of  
XX the breast or liver, Sezary syndrome and Gaucher's disease), neurological  
XX diseases (e.g. Alzheimer's disease, Parkinson's disease) cardio-  
XX /cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina),  
XX infections caused by bacteria, viruses and fungi and ocular disorders  
XX (e.g. corneal infections). Other uses include wound healing, maintenance  
XX of organs before transplantation, support of cell culture of primary

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178  
Db 47 DIVIVLDGNSIYPW 61

RESULT 48  
ASP47854  
ID ABP47854 standard; protein; 148 AA.  
XX  
XX AC ABP47854;  
XX  
XX DT 23-AUG-2002 (first entry)  
XX  
XX DE Human polypeptide SEQ ID NO 284.  
XX  
XX KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
XX immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnary;  
XX antiparkinsonian; antisticking; antianaemic; antiarthritic; cancer;  
XX antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
XX anti-allergic; antidiabetic; antiulcer; anticonvulsant; antifungal;  
XX antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
XX neurological disease; infection; nephrotropic; gene therapy; vaccine.  
XX  
XX OS Homo sapiens.  
XX  
XX PN US2002042386-A1.  
XX  
XX PD 11-APR-2002.  
XX  
XX PF 17-JAN-2001; 2001US-00764870.  
XX  
XX PR 31-JAN-2000; 2000US-0179065P.  
XX PR 04-FEB-2000; 2000US-0180628P.  
XX PR 28-JUN-2000; 2000US-0214886P.  
XX PR 07-JUL-2000; 2000US-0216647P.  
XX PR 11-JUL-2000; 2000US-0216880P.  
XX PR 11-JUL-2000; 2000US-0217487P.  
XX PR 11-JUL-2000; 2000US-0217496P.  
XX PR 14-JUL-2000; 2000US-0218290P.  
XX PR 26-JUL-2000; 2000US-0220963P.  
XX PR 14-AUG-2000; 2000US-0224518P.  
XX PR 14-AUG-2000; 2000US-0224519P.  
XX PR 14-AUG-2000; 2000US-0225267P.  
XX PR 14-AUG-2000; 2000US-0225268P.  
XX PR 14-AUG-2000; 2000US-0225270P.  
XX PR 14-AUG-2000; 2000US-0225447P.  
XX PR 14-AUG-2000; 2000US-0225757P.  
XX PR 22-AUG-2000; 2000US-0226868P.  
XX PR 30-AUG-2000; 2000US-0228924P.  
XX PR 01-SEP-2000; 2000US-0229287P.  
XX PR 01-SEP-2000; 2000US-0229344P.  
XX PR 01-SEP-2000; 2000US-0229345P.  
XX PR 05-SEP-2000; 2000US-0229509P.  
XX PR 05-SEP-2000; 2000US-0229513P.  
XX PR 08-SEP-2000; 2000US-0231413P.  
XX PR 21-SEP-2000; 2000US-0234223P.  
XX PR 21-SEP-2000; 2000US-0234274P.  
XX PR 25-SEP-2000; 2000US-0234997P.  
XX PR 27-SEP-2000; 2000US-0235834P.  
XX PR 29-SEP-2000; 2000US-0236327P.  
XX PR 29-SEP-2000; 2000US-0236367P.  
XX PR 29-SEP-2000; 2000US-0236368P.  
XX PR 29-SEP-2000; 2000US-0236369P.  
XX PR 29-SEP-2000; 2000US-0236370P.  
XX PR 02-OCT-2000; 2000US-0236802P.  
XX PR 02-OCT-2000; 2000US-0237037P.  
XX PR 02-OCT-2000; 2000US-0237038P.  
XX PR 02-OCT-2000; 2000US-0237039P.

Query Match 1.3%; Score 15; DB 4; Length 148;  
Best Local Similarity 100.0%; Pred. No. 6.9e-06;

PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239335P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0244817P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 XX (ROSE/) ROSEN C A.  
 PA (RUBE/) RUBEN S M.  
 PA (BARA/) BARASH S C.  
 XX  
 PI Rosen CA, Ruben SM, Barash SC;  
 DR  
 XX WPI; 2002-470713/50.  
 DR N-PSDB; ABQ66529.  
 XX  
 PT New nucleic acid encoding human proteins, useful for diagnosis, treatment  
 PT and prevention of e.g. osteoporosis, also related polypeptides and  
 PT antibodies.  
 XX  
 PS Claim 11; SEQ ID NO 284; 235pp + Sequence Listing; English.  
 CC  
 CC The invention relates to novel genes (ABQ66521-ABQ66785) and proteins  
 CC (ABP47846-ABP4810) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing  
 CC (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from USPTO at seqdata.uspto.gov/sequence.html?DocID=99909764870  
 XX  
 SQ Sequence 148 AA;  
 Query Match 1.3%; Score 15; DB 5; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 164 DIVIVLDGNSIYPW 178  
 Db 47 DIVIVLDGNSIYPW 61  
 RESULT 49  
 ID ABP48014 standard; protein; 148 AA.  
 XX  
 AC ABP48014;  
 XX  
 XX 23-AUG-2002 (first entry)  
 DT  
 XX Human polypeptide SEQ ID NO 444.  
 DE  
 XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnary;  
 KW antiparkinsonian; antiskinkling; antianaemic; antiarthritic; cancer;  
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
 KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;  
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine.  
 XX

OS Homo sapiens.  
 XX US2002042386-A1.  
 PN  
 XX  
 PD 11-APR-2002.  
 XX  
 PF 17-JAN-2001; 2001US-00764870.  
 XX  
 XX 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 22-AUG-2000; 2000US-0226888P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 XX  
 PA (ROSE/) ROSEN C A.  
 PA (RUBE/) RUBEN S M.  
 PA (BARA/) BARASH S C.  
 XX  
 PI Rosen CA, Ruben SM, Barash SC;  
 DR  
 XX WPI; 2002-470713/50.  
 DR N-PSDB; ABQ66589.  
 XX  
 PT New nucleic acid encoding human proteins, useful for diagnosis, treatment  
 PT and prevention of e.g. osteoporosis, also related polypeptides and  
 PT antibodies.  
 XX  
 PS Claim 11; SEQ ID NO 444; 235pp + Sequence Listing; English.





PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 06-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0253678P.  
PR 17-JAN-2001; 2001US-00764870.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM, Barash SC;  
XX  
XX WPI; 2003-743765/70.  
XX N-PSDB; ADC10551.  
XX  
XX New isolated nucleic acids and polypeptides, useful for diagnosing,  
XX treating, and/or preventing disorders, such as cancer, infections,  
XX cardiovascular and inflammatory diseases.  
XX  
XX Claim 11; SEQ ID NO 284; 235pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (cDNA)  
XX encoding a human extracellular matrix protein, representing one of 161  
XX novel genes. Also included are recombinant vectors, host cells  
XX (expressing the protein), the extracellular matrix proteins (including  
XX their fragments, epitopes and homologues), an isolated antibody that  
XX binds specifically to the protein, diagnosing a pathological condition or  
XX susceptibility to a pathological condition (comprising determining the  
XX presence or absence of a mutation in the nucleic acid and diagnosing a  
XX condition based on the presence or absence of the mutation), diagnosing a  
XX pathological condition or susceptibility to a pathological condition

CC (comprising determining the presence or amount of expression of the  
CC protein in a biological sample and diagnosing a condition based on the  
CC presence or amount of expression of the protein), preventing, treating or  
CC ameliorating a medical condition by administering a binding partner to the  
CC protein to a mammalian subject, identifying a binding partner to the  
CC protein, the gene corresponding to the cDNA sequence, and identifying an  
CC activity in a biological assay (comprising expressing the nucleic acid in  
CC a cell, isolating the supernatant, detecting an activity in a biological  
CC assay and identifying the protein in the supernatant having the  
CC activity). The nucleic acids and proteins display the following  
CC activities: Cytostatic, antibacterial, Virucide, Neuroprotective,  
CC Gynaecological, Gastrointestinal-Gen, Cardiant, Cardiovascular-Gen,  
CC Nephrotropic, Antiinflammatory, Muscular-Gen, Respiratory-Gen,  
Query Match 1.3%; Score 15; DB 7; Length 148;  
Best Local Similarity 100.0%; Pred. No. 6.9e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 164 DIVIVLDGNSIYPW 178  
DB 47 DIVIVLDGNSIYPW 61  
|||||  
|||||  
RESULT 51  
ADC10976  
ID ADC10976 standard; protein; 148 AA.  
XX  
XX AC ADC10976;  
XX  
XX DT 18-DEC-2003 (first entry)  
XX  
XX DE Human protein from extracellular matrix gene 9 #2.  
XX  
XX KW Extracellular matrix protein; cytostatic; antibacterial; virucide;  
KW neuroprotective; gynaecological; gastrointestinal-Gen; cardiant;  
KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;  
KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;  
KW neotropic; antiallergic; cancer; bacterial infection; viral infection;  
KW neural disorder; immune system disorder; blood disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW inflammatory disorder; proliferative disorder; Human.  
XX  
XX OS Homo sapiens.  
XX  
XX PN US2003059875-A1.  
XX  
XX PD 27-MAR-2003.  
XX  
XX PF 19-APR-2002; 2002US-00125540.  
XX  
XX PR 31-JAN-2000; 2000US-0179065P.  
XX PR 04-FEB-2000; 2000US-0180628P.  
XX PR 24-FEB-2000; 2000US-0184664P.  
XX PR 02-MAR-2000; 2000US-0186350P.  
XX PR 16-MAR-2000; 2000US-0189874P.  
XX PR 17-MAR-2000; 2000US-0190076P.  
XX PR 18-APR-2000; 2000US-0198123P.  
XX PR 19-MAY-2000; 2000US-0205515P.  
XX PR 07-JUN-2000; 2000US-0209467P.  
XX PR 28-JUN-2000; 2000US-0214886P.  
XX PR 30-JUN-2000; 2000US-0215135P.  
XX PR 07-JUL-2000; 2000US-0216647P.  
XX PR 07-JUL-2000; 2000US-0216880P.  
XX PR 11-JUL-2000; 2000US-0217487P.  
XX PR 11-JUL-2000; 2000US-0217496P.  
XX PR 14-JUL-2000; 2000US-0218290P.  
XX PR 26-JUL-2000; 2000US-0220963P.  
XX PR 26-JUL-2000; 2000US-0220964P.  
XX PR 14-AUG-2000; 2000US-0224518P.  
XX PR 14-AUG-2000; 2000US-0224519P.  
XX PR 14-AUG-2000; 2000US-0225213P.  
XX PR 14-AUG-2000; 2000US-0225214P.



CC a cell, isolating the supernatant, detecting an activity in a biological  
 CC assay and identifying the protein in the supernatant having the  
 CC activity). The nucleic acids and proteins display the following  
 CC activities Cytostatic, antibacterial, Virucide, Neuroprotective,  
 CC Gynaecological, Gastrointestinal-Gen, Cardiac, Cardiovascular-Gen,  
 CC Nephrologic, Antiinflammatory, Muscular-Gen, Respiratory-Gen,

Query Match 1.3%; Score 15; DB 7; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVIVLDGNSIYPW 178  
 Db 47 DIVIVLDGNSIYPW 61

RESULT 52  
 AAE33535  
 ID AAE33535 standard; protein; 192 AA.

XX AAE33535;

XX 23-OCT-2003 (revised)

DT 02-APR-2003 (first entry)

XX Rat-human alpha 1-I domain fusion protein, RdeltaH.

XX Human; very late activation antigen; VLA-1; betal containing integrin;  
 KW immunological disorder; inflammatory disorder; skin related condition;  
 KW psoriasis; eczema; burn; dermatitis; respiratory distress syndrome;  
 KW fibrosis; allergic rhinitis; asthma; bronchitis; tendonitis; bursitis;  
 KW fever; migraine headache; inflammatory bowel disease; Crohn's disease;  
 KW irritable bowel syndrome; colitis; colorectal cancer; vascular disease;  
 KW atherosclerosis; thyroiditis; aplastic anaemia; periarteritis nodosa;  
 KW gastritis; Hodgkin's disease; rheumatic fever; autoimmune disease;  
 KW osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis;  
 KW systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome;  
 KW renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis;  
 KW hypersensitivity; graft rejection; transplant rejection; conjunctivitis;  
 KW graft versus host disease; myocardial ischaemia; fusion protein; rat.

OS Homo sapiens.

OS Rattus sp.

OS Chimeric.

XX WO200283854-A2.

XX 24-OCT-2002.

XX 12-APR-2002; 2002WO-US011521.

XX 13-APR-2001; 2001US-0283794P.

XX 06-JUL-2001; 2001US-0303689P.

XX (BIOJ ) BIOGEN INC.

XX Lyne PD, Garber EA, Saldanha JW, Karpusas M;

XX WPI; 2003-093009/08.

XX New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-  
 PT mediated immunological or inflammatory disorders, e.g. psoriasis, eczema,  
 PT burns, dermatitis, and abnormal proliferation of hair follicle cells or  
 PT fibrosis.

XX Example 24; Page 93; 248pp; English.

XX The present invention relates to novel antibodies that specifically bind  
 CC to very late activation (VLA-1; betal containing integrins) antigens and  
 CC methods of using these antibodies to treat immunological disorders. The  
 CC anti-VLA-1 antibodies are useful for preventing or treating VLA-1-  
 CC mediated immunological or inflammatory disorders such as skin related  
 CC conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal

CC proliferation of hair follicle cells), fibrosis (e.g. kidney or lung  
 CC fibrosis), allergic rhinitis, respiratory distress syndrome, asthma,  
 CC bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-  
 CC intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease,  
 CC gastritis, irritable bowel syndrome, colitis and colorectal cancer),  
 CC vascular diseases (e.g. atherosclerosis), thyroiditis, aplastic anaemia,  
 CC periarteritis nodosa, Hodgkin's disease, rheumatic fever, osteoarthritis,  
 CC autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid  
 CC arthritis, systemic lupus erythematosus and multiple sclerosis), renal  
 CC failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis,  
 CC polymyositis, hypersensitivity (e.g. delayed type hypersensitivity or  
 CC immediate hypersensitivity), graft and transplant rejections, graft  
 CC versus host disease, conjunctivitis, swelling occurring after injury,  
 CC myocardial ischaemia or endotoxin shock syndrome. The present sequence is  
 CC rat-human alpha 1-I domain fusion protein, RdeltaH. This sequence is used  
 CC in the exemplification of the invention. (Updated on 23-OCT-2003 to  
 CC standardise OS field)

XX Sequence 192 AA;

Query Match 1.3%; Score 15; DB 6; Length 192;

Best Local Similarity 100.0%; Pred. No. 8.7e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVIVLDGNSIYPW 178

Db 4 DIVIVLDGNSIYPW 18

RESULT 53

AAE33536

ID AAE33536 standard; protein; 192 AA.

XX AAE33535;

XX 02-APR-2003 (first entry)

XX Rat alpha 1-I domain protein #1.

XX Rat; very late activation antigen; VLA-1; betal containing integrin;  
 KW immunological disorder; inflammatory disorder; skin related condition;  
 KW psoriasis; eczema; burn; dermatitis; respiratory distress syndrome;  
 KW fibrosis; allergic rhinitis; asthma; bronchitis; tendonitis; bursitis;  
 KW fever; migraine headache; inflammatory bowel disease; Crohn's disease;  
 KW irritable bowel syndrome; colitis; colorectal cancer; vascular disease;  
 KW atherosclerosis; thyroiditis; aplastic anaemia; periarteritis nodosa;  
 KW gastritis; Hodgkin's disease; rheumatic fever; autoimmune disease;  
 KW osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis;  
 KW systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome;  
 KW renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis;  
 KW hypersensitivity; graft rejection; transplant rejection; conjunctivitis;  
 KW graft versus host disease; myocardial ischaemia; alpha 1-I protein.

OS Rattus sp.

XX WO200283854-A2.

XX 24-OCT-2002.

XX 12-APR-2002; 2002WO-US011521.

XX 13-APR-2001; 2001US-0283794P.

XX 06-JUL-2001; 2001US-0303689P.

XX (BIOJ ) BIOGEN INC.

XX Lyne PD, Garber EA, Saldanha JW, Karpusas M;

XX WPI; 2003-093009/08.

XX New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-  
 PT mediated immunological or inflammatory disorders, e.g. psoriasis, eczema,  
 PT burns, dermatitis, and abnormal proliferation of hair follicle cells or

```

PT fibrosis.
PS Example 24; Page 93-94; 248pp; English.
XX
XX
CC The present invention relates to novel antibodies that specifically bind
CC to very late activation (VLA-1; beta1 containing integrins) antigens and
CC methods of using these antibodies to treat immunological disorders. The
CC anti-VLA-1 antibodies are useful for preventing or treating VLA-1-
CC mediated immunological or inflammatory disorders such as skin related
CC conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal
CC proliferation of hair follicle cells), fibrosis (e.g. kidney or lung
CC fibrosis), allergic rhinitis, respiratory distress syndrome, asthma,
CC bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-
CC intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease,
CC gastritis, irritable bowel syndrome, colitis and colorectal cancer),
CC vascular diseases (e.g. atherosclerosis), thyroiditis, aplastic anaemia,
CC periarthritis nodosa, Hodgkin's disease, rheumatic fever, osteoarthritis,
CC autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid
CC arthritis, systemic lupus erythematosus and multiple sclerosis), renal
CC failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis,
CC polynovositis, hypersensitivity (e.g. delayed type hypersensitivity or
CC immediate hypersensitivity), graft and transplant rejections, graft
CC versus host disease, conjunctivitis, swelling occurring after injury,
CC myocardial ischaemia or endotoxin shock syndrome. The present sequence is
CC rat 1-I domain protein. This sequence is used in the exemplification of
CC the invention
XX
SQ Sequence 192 AA;

Query Match 1.3%; Score 15; DB 6; Length 192;
Best Local Similarity 100.0%; Pred. No. 8.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
Db |||||
4 DIVIVLDGNSIYPW 18

RESULT 55
AAU76851
ID AAU76851 standard; protein; 195 AA.
XX
AC AAU76851;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human integrin alpha subunit Alpha 1 (CD49a) A domain.
XX
KW Human; integrin alpha subunit; A domain; CD49a; integrin beta subunit;
KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
KW ischaemia-reperfusion injury; immune complex; parasitic disease; Alpha 1;
KW antinflammatory; vasotropic; antiparasitic; vulnery; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200209737-A1.
XX
PD 07-FEB-2002.
XX
PF 31-JUL-2001; 2001WO-US023957.
XX
PR 31-JUL-2000; 2000US-0221950P.
PR 11-JAN-2001; 2001US-00758493.
PR 13-MAR-2001; 2001US-00805354.

```

```

XX PA (GEO) GEN HOSPITAL CORP.
XX PI
XX PI Annaout AM, Li R, Xiong J;
XX DR WPI; 2002-188687/24.
XX DR
XX PT Novel high affinity integrin polypeptide useful for treating restenosis
XX PT and parasitic diseases, comprises all or part of variant integrin alpha
XX PT subunit A domain or variant integrin beta subunit A-like domain.
XX PS
XX PS Example 2; Fig 5; 55pp; English.
XX CC The invention relates to a high affinity integrin polypeptide comprising
XX CC all or part of a variant integrin alpha subunit A domain or a variant
XX CC integrin beta subunit A-like domain. The polypeptide, preferably the
XX CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
XX CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
XX CC V at residue 315 and A at residue 320 have been replaced by C, is useful
XX CC for determining if a test compound is a candidate compound for binding to
XX CC CD11b or for treating an inflammatory disorder, by contacting a test
XX CC compound with the polypeptide and determining if the test compound binds
XX CC to the polypeptide. The integrin subunits are useful for reducing
XX CC skeletal muscle injury, for treating disorders caused by ischaemia-
XX CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
XX CC to purify variant integrin polypeptide ligands and as bait proteins in
XX CC two-hybrid or three-hybrid assays. This sequence represents the human
XX CC integrin alpha subunit Alpha 1 (CD49a) A domain
XX SQ Sequence 195 AA;
Query Match 1.3%; Score 15; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 164 DIVIVLDGNSIYPW 178
Db 6 DIVIVLDGNSIYPW 20
RESULT 56
AAU76860
ID AAU76860 standard; protein; 195 AA.
XX AC AAU76860;
XX DT 21-MAY-2002 (first entry)
XX DE Human integrin alpha subunit Alpha 1 (CD49a) variant A domain.
XX DE Human; integrin alpha subunit; A domain; CD49a; integrin beta subunit;
XX KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
XX KW ischaemia-reperfusion injury; immune complex; parasitic disease; mutant;
XX KW anti-inflammatory; vasotropic; antiparasitic; vulnerary; gene therapy;
XX KW mutein; Alpha 1.
XX OS Homo sapiens.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 193
XX FT /note= "Wild-type Ile substituted by any other amino
XX FT acid"
XX PN WO200209737-A1.
XX PD 07-FEB-2002.
XX PP 31-JUL-2001; 2001WO-US023957.
XX PR 31-JUL-2000; 2000US-0221950P.
XX PR 11-JAN-2001; 2001US-00758493.
XX PR 13-MAR-2001; 2001US-00805354.

```

```

XX PA (GEO) GEN HOSPITAL CORP.
XX PI
XX PI Annaout AM, Li R, Xiong J;
XX DR WPI; 2002-188687/24.
XX DR
XX PT Novel high affinity integrin polypeptide useful for treating restenosis
XX PT and parasitic diseases, comprises all or part of variant integrin alpha
XX PT subunit A domain or variant integrin beta subunit A-like domain.
XX PS
XX PS Claim 45; Page; 55pp; English.
XX CC The invention relates to a high affinity integrin polypeptide comprising
XX CC all or part of a variant integrin alpha subunit A domain or a variant
XX CC integrin beta subunit A-like domain. The polypeptide, preferably the
XX CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
XX CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
XX CC V at residue 315 and A at residue 320 have been replaced by C, is useful
XX CC for determining if a test compound is a candidate compound for binding to
XX CC CD11b or for treating an inflammatory disorder, by contacting a test
XX CC compound with the polypeptide and determining if the test compound binds
XX CC to the polypeptide. The integrin subunits are useful for reducing
XX CC skeletal muscle injury, for treating disorders caused by ischaemia-
XX CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
XX CC to purify variant integrin polypeptide ligands and as bait proteins in
XX CC two-hybrid or three-hybrid assays. This sequence represents a human
XX CC integrin alpha subunit Alpha 1 (CD49a) variant A domain. Note: This
XX CC variant sequence is not featured in the specification but has been
XX CC derived from the wild-type protein shown in AAU76851
XX SQ Sequence 195 AA;
Query Match 1.3%; Score 15; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 164 DIVIVLDGNSIYPW 178
Db 6 DIVIVLDGNSIYPW 20
RESULT 57
AAU76860
ID AAU76860 standard; protein; 214 AA.
XX AC AAU76860;
XX DT 14-MAR-2001 (first entry)
XX DE Rat alpha1 integrin domain.
XX DE Rat; alpha1 integrin domain; alphasbeta1 function blocking antibody;
XX KW inflammatory disorder; rheumatoid arthritis; skin condition; asthma;
XX KW bronchitis; headache; antipyretic; fever; gastrointestinal;
XX KW vascular disease; autoimmune; respiratory distress syndrome;
XX KW endotoxin shock syndrome; atherosclerosis.
XX OS Rattus sp.
XX PN WO2000072881-A1.
XX PD 07-DEC-2000.
XX PP 01-JUN-2000; 2000WO-US015004.
XX PR 01-JUN-1999; 99US-0137038P.
XX PR 29-FEB-2000; 2000US-0185336P.
XX PA (BIO) BIOGEN INC.
XX PI De Fougereolles A, Gotwals P, Lobb R, Koteliarsky V;
XX

```

DR WPI; 2001-061448/07.  
 XX Use of blocking monoclonal antibody capable of binding to an epitope of  
 PT VLA-1 for treating inflammatory disorders, in particular arthritis.  
 XX  
 PS Example 14; Fig 14; 60pp; English.  
 XX  
 CC The present invention relates to the use of an alphanbetal function  
 CC blocking antibody capable of binding an epitope of human alphan integrin  
 CC domain (VLA-1; see AAB50043). The antibody of the present invention is  
 CC useful for treating an inflammatory disorder e.g. rheumatoid arthritis,  
 CC skin related conditions such as psoriasis, eczema, burns and dermatitis,  
 CC asthma, bronchitis, menstrual cramps, tendinitis, bursitis, and the  
 CC treatment of pain and headaches, or as an antipyretic, for the treatment  
 CC of fever, gastrointestinal conditions such as inflammatory bowel  
 CC diseases, Crohn's disease, gastritis and vascular diseases, migraine  
 CC headaches, periarthritis nodosa, thyroiditis, aplastic anaemia, Hodgkin's  
 CC disease, rheumatic fever, type I diabetes, myasthenia gravis, multiple  
 CC sclerosis, sarcoidosis, nephrotic syndrome, myocardial ischaemia,  
 CC allergic rhinitis, respiratory distress syndrome, endotoxin shock  
 CC syndrome and atherosclerosis. The present sequence is the rat alphan  
 CC integrin domain  
 XX  
 SQ Sequence 214 AA;  
 Query Match 1.3%; Score 15; DB 4; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 164 DIVIVLDGNSIYPW 178  
 |||||  
 Db 22 DIVIVLDGNSIYPW 36  
 |||||  
 RESULT 58  
 AAB50042  
 ID AAB50042 standard; protein; 214 AA.  
 AC AAB50042;  
 XX  
 DT 14-MAR-2001 (first entry)  
 DE Human alphan integrin domain.  
 XX  
 KW Human; alphan integrin domain; alphanbetal function blocking antibody;  
 KW inflammatory disorder; rheumatoid arthritis; skin condition; asthma;  
 KW bronchitis; headache; antipyretic; fever; gastrointestinal;  
 KW vascular disease; autoimmune; respiratory distress syndrome;  
 KW endotoxin shock syndrome; atherosclerosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200072881-A1.  
 XX  
 PD 07-DEC-2000.  
 XX  
 PF 01-JUN-2000; 2000WO-US015004.  
 XX  
 PR 01-JUN-1999; 99US-0137038P.  
 PR 29-FEB-2000; 2000US-0185336P.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 XX  
 PI De Fougereolles A, Gotwals P, Lobb R, Koteliansky V;  
 XX  
 DR WPI; 2001-061448/07.  
 XX  
 CC Use of blocking monoclonal antibody capable of binding to an epitope of  
 PT VLA-1 for treating inflammatory disorders, in particular arthritis.  
 XX  
 PS Example 14; Fig 15; 60pp; English.  
 XX  
 CC The present invention relates to the use of an alphanbetal function

CC blocking antibody capable of binding an epitope of human alphan integrin  
 CC domain (VLA-1; see AAB50043). The antibody of the present invention is  
 CC useful for treating an inflammatory disorder e.g. rheumatoid arthritis,  
 CC skin related conditions such as psoriasis, eczema, burns and dermatitis,  
 CC asthma, bronchitis, menstrual cramps, tendinitis, bursitis, and the  
 CC treatment of pain and headaches, or as an antipyretic, for the treatment  
 CC of fever, gastrointestinal conditions such as inflammatory bowel  
 CC diseases, Crohn's disease, gastritis and vascular diseases, migraine  
 CC headaches, periarthritis nodosa, thyroiditis, aplastic anaemia, Hodgkin's  
 CC disease, rheumatic fever, type I diabetes, myasthenia gravis, multiple  
 CC sclerosis, sarcoidosis, nephrotic syndrome, myocardial ischaemia,  
 CC allergic rhinitis, respiratory distress syndrome, endotoxin shock  
 CC syndrome and atherosclerosis. The present sequence is the human alphan  
 CC integrin domain  
 XX  
 SQ Sequence 214 AA;  
 Query Match 1.3%; Score 15; DB 4; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 164 DIVIVLDGNSIYPW 178  
 |||||  
 Db 22 DIVIVLDGNSIYPW 36  
 |||||  
 RESULT 59  
 AAE33539  
 ID AAE33539 standard; protein; 214 AA.  
 AC AAE33539;  
 XX  
 DT 02-APR-2003 (first entry)  
 DE Rat alphan 1-1 domain protein #2.  
 XX  
 KW Rat; very late activation antigen; VLA-1; betal containing integrin;  
 KW immunological disorder; inflammatory disorder; skin related condition;  
 KW psoriasis; eczema; burn; dermatitis; respiratory distress syndrome;  
 KW fibrosis; allergic rhinitis; asthma; bronchitis; tendinitis; bursitis;  
 KW fever; migraine headache; inflammatory bowel disease; Crohn's disease;  
 KW irritable bowel syndrome; colitis; colorectal cancer; vascular disease;  
 KW atherosclerosis; thyroiditis; aplastic anaemia; periarthritis nodosa;  
 KW gastritis; Hodgkin's disease; rheumatic fever; autoimmune disease;  
 KW osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis;  
 KW systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome;  
 KW renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis;  
 KW hypersensitivity; graft rejection; transplant rejection; conjunctivitis;  
 KW graft versus host disease; myocardial ischaemia; alpha 1-1 protein.  
 XX  
 OS Rattus sp.  
 XX  
 PN WO200283854-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 12-APR-2002; 2002WO-US011521.  
 XX  
 PR 13-APR-2001; 2001US-0283794P.  
 PR 06-JUL-2001; 2001US-0303689P.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 XX  
 PI Lyne PD, Garber EA, Saldanha JW, Karpusas M;  
 XX  
 DR WPI; 2003-093009/08.  
 XX  
 PT New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-  
 PT mediated immunological or inflammatory disorders, e.g. psoriasis, eczema,  
 PT burns, dermatitis, and abnormal proliferation of hair follicle cells or  
 PT fibrosis.  
 XX  
 PS Example 18; Fig 11A; 248pp; English.

XX The present invention relates to novel antibodies that specifically bind  
 CC to very late activation (VLA-1; beta1 containing integrins) antigens and  
 CC methods of using these antibodies to treat immunological disorders. The  
 CC anti-VLA-1 antibodies are useful for preventing or treating VLA-1-  
 CC mediated immunological or inflammatory disorders such as skin related  
 CC conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal  
 CC proliferation of hair follicle cells), fibrosis (e.g. kidney or lung  
 CC fibrosis), allergic rhinitis, respiratory distress syndrome, asthma,  
 CC bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-  
 CC intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease,  
 CC gastritis, irritable bowel syndrome, colitis and colorectal cancer),  
 CC vascular diseases (e.g. atherosclerosis), thyroiditis, aplastic anaemia,  
 CC periarthritis nodosa, Hodgkin's disease, rheumatic fever, osteoarthritis,  
 CC autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid  
 CC arthritis, systemic lupus erythematosus and multiple sclerosis), renal  
 CC failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis,  
 CC polymyositis, hypersensitivity (e.g. delayed type hypersensitivity or  
 CC immediate hypersensitivity), graft and transplant rejections, graft  
 CC versus host disease, conjunctivitis, swelling occurring after injury,  
 CC myocardial ischaemia or endotoxin shock syndrome. The present sequence is  
 CC rat 1-I domain protein. This sequence is used in the exemplification of  
 CC the invention  
 CC  
 CC Sequence 214 AA;

Query Match 1.3%; Score 15; DB 6; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 164 DIVVLGDSNSIYPW 178  
 DB 22 DIVVLGDSNSIYPW 36  
 |||||

RESULT 60  
 AAE33540  
 ID AAE33540 standard; protein; 214 AA.  
 XX AC AAE33540;  
 XX DT 02-APR-2003 (first entry)  
 XX DE Human alpha 1-I domain protein #2.  
 XX KW Human; very late activation antigen; VLA-1; beta1 containing integrin;  
 KW immunological disorder; inflammatory disorder; skin related condition;  
 KW psoriasis; eczema; burn; dermatitis; respiratory distress syndrome;  
 KW fibrosis; allergic rhinitis; asthma; bronchitis; tendonitis; bursitis;  
 KW fever; migraine headache; inflammatory bowel disease; Crohn's disease;  
 KW irritable bowel syndrome; colitis; colorectal cancer; vascular disease;  
 KW atherosclerosis; thyroiditis; aplastic anaemia; periarthritis nodosa;  
 KW gastritis; Hodgkin's disease; rheumatic fever; autoimmune disease;  
 KW osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis;  
 KW systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome;  
 KW renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis;  
 KW hypersensitivity; graft rejection; transplant rejection; conjunctivitis;  
 KW graft versus host disease; myocardial ischaemia; alpha 1-I protein.  
 XX OS Homo sapiens.  
 XX PN WO200283854-A2.  
 XX PD 24-OCT-2002.  
 XX FF 12-APR-2002; 2002WO-US011521.  
 XX PR 13-APR-2001; 2001US-0283794P.  
 XX PR 06-JUL-2001; 2001US-0303689P.  
 XX XX (BIOJ) BIOGEN INC.  
 XX PA Lyne PD, Garber EA, Saldanha JW, Karpusas M;  
 XX PI

XX WPI; 2003-093009/08.  
 XX New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-  
 CC mediated immunological or inflammatory disorders, e.g. psoriasis, eczema,  
 CC burns, dermatitis, and abnormal proliferation of hair follicle cells or  
 CC fibrosis.

Example 18; Fig 12; 248pp; English.

XX The present invention relates to novel antibodies that specifically bind  
 CC to very late activation (VLA-1; beta1 containing integrins) antigens and  
 CC methods of using these antibodies to treat immunological disorders. The  
 CC anti-VLA-1 antibodies are useful for preventing or treating VLA-1-  
 CC mediated immunological or inflammatory disorders such as skin related  
 CC conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal  
 CC proliferation of hair follicle cells), fibrosis (e.g. kidney or lung  
 CC fibrosis), allergic rhinitis, respiratory distress syndrome, asthma,  
 CC bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-  
 CC intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease,  
 CC gastritis, irritable bowel syndrome, colitis and colorectal cancer),  
 CC vascular diseases (e.g. atherosclerosis), thyroiditis, aplastic anaemia,  
 CC periarthritis nodosa, Hodgkin's disease, rheumatic fever, osteoarthritis,  
 CC autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid  
 CC arthritis, systemic lupus erythematosus and multiple sclerosis), renal  
 CC failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis,  
 CC polymyositis, hypersensitivity (e.g. delayed type hypersensitivity or  
 CC immediate hypersensitivity), graft and transplant rejections, graft  
 CC versus host disease, conjunctivitis, swelling occurring after injury,  
 CC myocardial ischaemia or endotoxin shock syndrome. The present sequence is  
 CC human 1-I domain protein. This sequence is used in the exemplification of  
 CC the invention  
 CC  
 CC Sequence 214 AA;

Query Match 1.3%; Score 15; DB 6; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 164 DIVVLGDSNSIYPW 178  
 DB 22 DIVVLGDSNSIYPW 36  
 |||||

RESULT 61  
 ADE86652  
 ID ADE86652 standard; protein; 1151 AA.  
 XX AC ADE86652;  
 XX DT 29-JAN-2004 (first entry)  
 XX DE Human integrin alpha 1 subunit.  
 XX KW human; secreted protein; cancer; liver disorder; hepatitis;  
 KW neural disorder; Alzheimer's disease; integrin alpha 1 subunit.  
 XX OS Homo sapiens.  
 XX PN US2003129685-A1.  
 XX PD 10-JUL-2003.  
 XX FF 18-APR-2001; 2001US-00836353.  
 XX PR 28-OCT-1998; 98US-0105971P.  
 XX PR 27-OCT-1999; 99WO-US025031.  
 XX PR 19-APR-2000; 2000US-0198407P.  
 XX XX (NIJ) NI J.  
 XX PA (YOUNG) YOUNG P E.  
 XX PA (KENN) KENNY J J.  
 XX PA (OLSEN) OLSEN H S.



PA (MOORE/) MOORE P A.  
PA (WEI/) WEI Y.  
PA (GREENE/) GREENE J M.  
PA (RUBEN/) RUBEN S M.  
XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JW;  
PI Ruben SM;  
XX WPI; 2004-020335/02.  
XX New nucleic acid molecule, useful for preparing a medicament for  
PT preventing, treating or ameliorating a medical condition e.g. cancer,  
PT liver disorders or neural disorders.  
XX Disclosure; SEQ ID NO 103; 380pp; English.  
XX The invention relates to an isolated nucleic acid sequence, or its  
CC allelic variant, a fragment of the cDNA sequence, or its fragment,  
CC domain, epitope or species homologue. The nucleic acid is useful for  
CC preparing a medicament for preventing, treating or ameliorating a medical  
CC condition e.g., cancer, liver disorders such as hepatitis or neural  
CC disorders such as Alzheimer's disease. The present sequence represents  
CC the amino acid sequence of an integrin alpha 1 subunit.  
XX Sequence 1151 AA;  
SQ

Query Match 1.3%; Score 15; DB 8; Length 1151;  
Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 164 DIVIVLDGNSIYPW 178  
DB 144 DIVIVLDGNSIYPW 158  
|||||

RESULT 62  
ABB90759  
ID ABB90759 standard; protein; 1179 AA.  
AC ABB90759;  
XX 30-MAY-2002 (first entry)  
XX Human Tumour Endothelial Marker polypeptide SEQ ID NO 250.  
XX Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;  
XX normal endothelial marker; pan-endothelial marker; immunostimulant;  
XX antiangiogenic; tumour; neoangiogenesis; vascularised tumour;  
XX polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;  
XX psoriasis.  
XX Homo sapiens.  
XX WO200210217-A2.  
XX 07-FEB-2002.  
XX 01-AUG-2001; 2001WO-US024031.  
XX 02-AUG-2000; 2000US-0222599P.  
XX 11-AUG-2000; 2000US-0224360P.  
XX 11-APR-2001; 2001US-0282850P.  
XX (UYUO ) UNIV JOHNS HOPKINS.  
XX St Croix B, Kinzler KW, Vogelstein B;  
XX WPI; 2002-291856/33.  
XX N-PSDB; ABL92113.  
XX An isolated molecule comprising an antibody variable region which  
PT specifically binds to an extracellular domain of a tumor endothelial  
PT marker (TEM) protein, useful for inhibiting tumor growth.

XX Claim 35; Page 242-245; 331pp; English.  
XX The invention relates to an isolated molecule comprising an antibody  
CC variable region which specifically binds to an extracellular domain of a  
CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,  
CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM  
CC proteins have cytostatic, immunostimulant and antiangiogenic activity.  
CC They are useful for inhibiting tumour growth, neoangiogenesis in subjects  
CC bearing a vascularised tumour, polycystic kidney disease, diabetic  
CC retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM  
CC genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789)  
CC are disclosed, as are marker oligonucleotide sequences: tumour  
CC endothelial markers (TEM) ABL91936-ABL92041 and ABL92143-ABL92191; normal  
CC endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers  
CC (PEM) ABL91903-ABL91995  
XX Sequence 1179 AA;  
SQ

Query Match 1.3%; Score 15; DB 5; Length 1179;  
Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 164 DIVIVLDGNSIYPW 178  
DB 172 DIVIVLDGNSIYPW 186  
|||||

RESULT 63  
ABP64915  
ID ABP64915 standard; protein; 1179 AA.  
XX ABP64915;  
XX 25-FEB-2003 (first entry)  
XX Human protein SEQ ID 575.  
XX Human; expressed sequence tag; EST; haematopoietic disorder;  
XX central nervous system disease; viral infection;  
XX peripheral nervous system disease; non-healing wound; infectious disease;  
XX immune deficiency; immune disorder; bacterial infection; allergy; cancer;  
XX fungal infection; autoimmune disorder; coagulation disorder; neutropenic;  
XX anti-allergic; anti-inflammatory; immunosuppressive; neuroprotective;  
XX cytostatic; haemostatic; virucide; antibacterial; fungicide;  
XX immunostimulant; cerebroprotective.  
XX Homo sapiens.  
XX WO200259260-A2.  
XX 01-AUG-2002.  
XX 16-NOV-2001; 2001WO-US042950.  
XX 17-NOV-2000; 2000US-00714936.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;  
XX Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX WPI; 2002-590824/63.  
XX N-PSDB; ABQ99501.  
XX New isolated polynucleotide, useful in research, diagnostic or  
PT therapeutic methods, e.g. preventing or treating disorders involving  
PT aberrant protein expression or biological activity.  
XX Claim 20; SEQ ID NO 575; 394pp; English.  
XX The present invention relates to novel human coding sequences (ABQ99268-  
CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in

CC therapeutic, diagnostic and research methods. The polynucleotides may be  
 CC used in the field of molecular biology as hybridisation probes, primers  
 CC for PCR, for chromosome and gene mapping, for the recombinant production  
 CC of protein, or in generation of anti-sense DNA or RNA. The  
 CC polynucleotides are useful in diagnostics as expressed sequence tags  
 CC (ESTs) for identifying expressed genes or for physical mapping of the  
 CC human genome. The proteins may be used as molecular weight markers, or as  
 CC nutritional sources or supplements. The proteins may be used to maintain  
 CC and expand cell population in a totipotent or pluripotent state  
 CC useful for re-engineering damaged or diseased tissues, transplantation.  
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The  
 CC polynucleotides and proteins are useful for preventing, treating or  
 CC ameliorating disorders involving aberrant protein expression or  
 CC biological activity, e.g. haematopoietic disorders, central/peripheral  
 CC nervous system diseases, mechanical and traumatic disorders, non-healing  
 CC wounds, immune deficiencies and disorders, infectious diseases caused by  
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic  
 CC reactions and conditions, coagulation disorders, or cancer. The  
 CC polynucleotide sequences of the invention were assembled from ESTs  
 CC isolated mainly by sequencing by hybridisation, and in some cases,  
 CC sequences obtained from one or more public databases. Note: The sequence  
 CC data for this patent did not form part of the printed specification, but  
 CC was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1179 AA;

Query Match 1.3%; Score 15; DB 5; Length 1179;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q/ 164 DIVIVLDGNSIYPW 178  
 Db 172 DIVIVLDGNSIYPW 186  
 |||||

RESULT 64  
 ABUS4466  
 ID ABUS4466 standard; protein; 1179 AA.  
 AC ABUS4466;  
 XX  
 XX 12-MAR-2003 (first entry)  
 DT  
 XX Human tumour endothelial marker TEM 30.  
 DE  
 XX Human; endothelial cell; EC; tumour endothelial cell; TEM; NEM;  
 KW Tumour endothelial marker; normal endothelial marker; PEM;  
 KW pan-endothelial marker; polycystic kidney disease; psoriasis;  
 KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;  
 KW neovascularization; immune response; cytostatic; antidiabetic;  
 KW ophthalmological; antirheumatic; antiarthritic; antipsoriatic.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200283874-A2.  
 PN  
 XX 24-OCT-2002.  
 PD  
 XX 10-APR-2002; 2002WO-US008253.  
 PF  
 XX 11-APR-2001; 2001US-0282850P.  
 PR 06-FEB-2002; 2002US-0354262P.  
 XX  
 XX (UYJO ) UNIV JOHNS HOPKINS.  
 PA  
 XX Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;  
 PI  
 XX WPI: 2003-093016/08.  
 DR N-PSDB; AEX72038.  
 XX

PT New purified human transmembrane protein, designated as tumor endothelial  
 marker (TEM) 3, useful for detecting, diagnosing or treating tumors,  
 XX

PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or  
 PT psoriasis.  
 XX  
 PS Disclosure; Page 267-269; 374pp; English.  
 XX  
 CC The present invention relates to a novel method for the isolation of  
 CC endothelial cells (ECs), and the identification of genes expressed in  
 CC normal and tumour ECs. Tumour endothelial marker (TEM), normal  
 CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are  
 CC identified in human ECs. The human EC marker proteins and the  
 CC polynucleotide sequences encoding them are useful for detecting,  
 CC diagnosing or treating tumours as well as polycystic kidney disease,  
 CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also  
 CC useful for inhibiting neovascularization or tumour angiogenesis, for  
 CC inducing an immune response to tumour endothelial cells in a patient, or  
 CC for identifying candidate drugs for treating tumours. The present  
 CC sequence represents a human TEM or NEM protein of the invention  
 XX  
 SQ Sequence 1179 AA;

Query Match 1.3%; Score 15; DB 6; Length 1179;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q/ 164 DIVIVLDGNSIYPW 178  
 Db 172 DIVIVLDGNSIYPW 186  
 |||||

RESULT 65  
 ABB90788  
 ID ABB90788 standard; protein; 1180 AA.  
 XX  
 AC ABB90788;  
 XX  
 XX 30-MAY-2002 (first entry)  
 DT  
 XX Rat Tumour Endothelial Marker polypeptide SEQ ID NO 307.  
 DE  
 XX Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;  
 KW normal endothelial marker; pan-endothelial marker; immunostimulant;  
 KW antiangiogenic; tumour; neovascularization; vascularised tumour;  
 KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;  
 KW psoriasis.  
 XX  
 XX Rattus sp.  
 OS  
 XX WO200210217-A2.  
 PN  
 XX 07-FEB-2002.  
 PD  
 XX 01-AUG-2001; 2001WO-US024031.  
 PF  
 XX 02-AUG-2000; 2000US-0222599P.  
 PR 11-AUG-2000; 2000US-0224360P.  
 PR 11-APR-2001; 2001US-0282850P.  
 XX  
 XX (UYJO ) UNIV JOHNS HOPKINS.  
 PA  
 XX St Croix B, Kinzler KW, Vogelstein B;  
 PI  
 XX WPI: 2002-291856/33.  
 DR N-PSDB; ABL92141.  
 XX

PT An isolated molecule comprising an antibody variable region which  
 PT specifically binds to an extracellular domain of a tumor endothelial  
 PT marker (TEM) protein, useful for inhibiting tumor growth.  
 XX  
 PS Disclosure; Page 314-318; 331pp; English.  
 XX  
 CC The invention relates to an isolated molecule comprising an antibody  
 CC variable region which specifically binds to an extracellular domain of a  
 CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,  
 CC

CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM  
 CC proteins have cytostatic, immunostimulant and antiangiogenic activity.  
 CC They are useful for inhibiting tumour growth, neoangiogenesis in subjects  
 CC bearing a vascularised tumour, polycystic kidney disease, diabetic  
 CC retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM  
 CC genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789)  
 CC are disclosed, as are marker oligonucleotide sequences: tumour  
 CC endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal  
 CC endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers  
 CC (PEM) ABL91903-ABL91995

XX SQ Sequence 1180 AA;  
 Query Match 1.3%; Score 15; DB 5; Length 1180;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVIVLDGNSIYPW 178  
 Db 172 DIVIVLDGNSIYPW 186  
 |||||

RESULT 66  
 ABB90749  
 ID ABB90749 standard; protein; 1180 AA.

XX AC ABB90749;  
 XX DT 12-MAR-2003 (first entry)  
 XX DE Mouse tumour endothelial marker mTEM 30.

XX KW Mouse; endothelial cell; EC; tumour endothelial cell; TEM; NEM;  
 KW Tumour endothelial marker; normal endothelial marker; PEM;  
 KW pan-endothelial marker; polycystic kidney disease; psoriasis;  
 KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;  
 KW neoangiogenesis; immune response; cytostatic; antidiabetic;  
 KW ophthalmological; antirheumatic; antiarthritic; antipsoriatic.

XX OS Mus sp.

XX PN WO200283874-A2.

XX PD 24-OCT-2002.

XX PF 10-APR-2002; 2002WO-US008253.

XX PR 11-APR-2001; 2001US-0282850P.

XX PR 06-FEB-2002; 2002US-0354262P.

XX PA (UNJO ) UNIV JOHNS HOPKINS.

XX PI Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;

XX DR WPI; 2003-093016/08.

XX DR N-PSDB; ABX72066.

XX PT New purified human transmembrane protein, designated as tumor endothelial  
 PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors,  
 PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or  
 PT psoriasis.

XX PS Disclosure; Page 356-358; 374pp; English.

XX CC The present invention relates to a novel method for the isolation of  
 CC endothelial cells (ECs), and the identification of genes expressed in  
 CC normal and tumour ECs. Tumour endothelial marker (TEM), normal  
 CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are  
 CC identified in human ECs. The human EC marker proteins and the  
 CC polynucleotide sequences encoding them are useful for detecting,  
 CC diagnosing or treating tumours as well as polycystic kidney disease,  
 CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also  
 CC useful for inhibiting neoangiogenesis or tumour angiogenesis, for

CC inducing an immune response to tumour endothelial cells in a patient, or  
 CC for identifying candidate drugs for treating tumours. The present  
 CC sequence represents a mouse TEM protein

XX SQ Sequence 1180 AA;

Query Match 1.3%; Score 15; DB 6; Length 1180;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVIVLDGNSIYPW 178  
 Db 172 DIVIVLDGNSIYPW 186  
 |||||

RESULT 67  
 ABB90750  
 ID ABB90750 standard; protein; 1180 AA.

XX AC ABB90750;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein P18614, SEQ ID NO 9512.

XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX PN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEHO ) GEN HOSPITAL CORP.

XX PA (FARB ) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX DR WPI; 2003-288312/26.

XX DR GENBANK; P18614.

XX PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating

CC pain (e.g. spinal segmental nerve injury (Chung)), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1180 AA;

Query Match 1.3%; Score 15; DB 7; Length 1180;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYFW 178  
 Db 172 DIVVLGSGNSIYFW 186

# RESULT 68

AAB30928  
 ID AAB30928 standard; peptide; 15 AA.

XX AC AAB30928;

XX DT 02-APR-2001 (first entry)

XX DE Antigenic peptide derived from a human alpha11 integrin chain.

XX KW Human; integrin; alpha11 subunit; fibroblast; muscle cell; chondrocyte;  
 KW osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;  
 KW wound healing; trauma; rheumatoid arthritis; osteoarthritis;  
 KW osteoporosis; cartilage damage; bone damage; cartilage.

XX OS Homo sapiens.

XX FN WO200075197-A1.

XX PD 14-DEC-2000.

XX PP 31-MAY-2000; 2000WO-SE001135.

XX FR 03-JUN-1999; 99SE-00002056.

XX PA (ACT1-) ACTIVE BIOTECH AB.

XX PI Gullberg D;

XX DR WPI; 2001-071061/08.

XX PT Integrin subunit alpha 11 or integrin heterodimer comprising subunit  
 PT alpha 11 in association with subunit beta, useful for treating muscle  
 PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.

XX PS Disclosure; Page 17; 79pp; English.

XX CC The present sequence is derived from the cytoplasmic domain of the human  
 CC integrin subunit, designated alpha11. The alpha11 polynucleotide and  
 CC polypeptide are useful as markers of cell target molecules, such as  
 CC fibroblasts, muscle cells, chondrocytes, osteoblasts, mesenchymally  
 CC derived cells or stem cells. They are also used for determining the  
 CC differential stage of cells during differentiation, development in  
 CC pathological conditions, in tissue regeneration, in transplantation or in  
 CC therapeutic and physiological repair of tissues. The pathological  
 CC conditions involving subunit alpha11 are selected from damage of cells,  
 CC muscle dystrophy, fibrosis, wound healing, trauma, rheumatoid arthritis,  
 CC osteoarthritis and osteoporosis, damage of cartilage and bone, and  
 CC cartilage and bone diseases. The polypeptide is useful for detecting the  
 CC formation of cartilage during embryonic development, for detecting  
 CC physiological therapeutic repair of cartilage and muscle, for selection  
 CC and analysis, or for sorting, isolating or purification of chondrocytes  
 CC and muscle cells, for detecting regeneration of cartilage or chondrocytes  
 CC during transplantation of cartilage or chondrocytes during

CC transplantation of cartilage or chondrocytes, respectively, or of muscle  
 CC or muscle cells during transplantation of muscle or muscle cells,  
 CC respectively, and for studies of differentiation or chondrocytes or  
 CC muscle cells

XX Sequence 15 AA;

Query Match 1.2%; Score 14; DB 4; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1175 REEPGLDTPKYLE 1188

Db 2 REEPGLDTPKYLE 15

# RESULT 69

AAV07728  
 ID AAV07728 standard; protein; 1183 AA.

XX AC AAV07728;

XX DT 01-JUL-1999 (first entry)

XX DE Armenian hamster alpha-1 integrin subunit protein.

XX KW VEGF; tumour angiogenesis inhibition; vascular endothelial growth factor;  
 KW integrin cell surface receptor; capillary; blood vessel; hamster;  
 KW alpha-1 subunit; alpha-2 subunit.

XX OS Cricetulus migratorius.

XX PN WO9916465-A1.

XX PD 08-APR-1999.

XX PP 30-SEP-1997; 97WO-US017485.

XX PR 30-SEP-1997; 97WO-US017485.

XX PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX PA (SENG/) SENG D R.

XX PA (DETW/) DETMAR M.

XX PA (CLAF/) CLAFFEY K P.

XX PI Senger DR, Detmar M, Claffey KP;

XX DR WPI; 1999-254930/21.

XX PT Inhibition of tumor angiogenesis through interaction of vascular  
 PT endothelial growth factor and integrin cell surface receptors.

XX PS Disclosure; Fig 2A-C; 64pp; English.

XX CC This invention describes a novel method for the inhibition of tumour  
 CC angiogenesis mediated by vascular endothelial growth factor (VEGF) and  
 CC integrin cell surface receptors expressed in vasculature of living  
 CC subjects. The method inhibits new capillary and new blood vessel  
 CC formation both within a tumour mass itself as well as in the immediately  
 CC adjacent blood vasculature surrounding the perimeter of the tumour mass.  
 CC Interaction and dependence upon VEGF to induce specific integrin  
 CC heterodimers in tumour angiogenesis provides a novel method for  
 CC inhibiting tumour angiogenesis, and unlike prior art relies on the  
 CC specific inter-relationship of VEGF and integrins, rather than  
 CC concentrating solely on one specific class of protein

XX Sequence 1183 AA;

Query Match 1.2%; Score 14; DB 2; Length 1183;  
 Best Local Similarity 100.0%; Pred. No. 0.00042;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYP 177

```

Db      174 DIVLDGNSIYP 187
|||||
RESULT 70
AAU76862
ID AAU76862 standard; protein; 195 AA.
XX AC AAU76862;
XX DT 21-MAY-2002 (first entry)
XX DE Human integrin alpha subunit Alpha 10 variant A domain.
XX KW Human; integrin alpha subunit; A domain; Alpha 10; integrin beta subunit;
XX KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
XX KW ischaemia-reperfusion injury; immune complex; parasitic disease; mutant;
XX KW antiinflammatory; vasotropic; antiparasitic; vulnery; gene therapy;
XX KW muten.
XX OS Homo sapiens.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Misc-difference 193
FT FT /note= "Wild-type Ile substituted by any other amino
FT FT acid"
XX PN WO200209737-A1.
XX XX 07-FEB-2002.
XX XX 31-JUL-2001; 2001WO-US023957.
XX PR 31-JUL-2000; 2000US-0221950P.
XX PR 11-JAN-2001; 2001US-00758493.
XX PR 13-MAR-2001; 2001US-00805354.
XX PA (GHEO ) GEN HOSPITAL CORP.
XX PI Arnaout AM, Li R, Xiong J;
XX DR WPI; 2002-188687/24.
XX PT Novel high affinity integrin polypeptide useful for treating restenosis
XX PT and parasitic diseases, comprises all or part of variant integrin alpha
XX PT subunit A domain or variant integrin beta subunit A-like domain.
XX PS Claim 53; Page; 55pp; English.
XX CC The invention relates to a high affinity integrin polypeptide comprising
XX CC all or part of a variant integrin alpha subunit A domain or a variant
XX CC integrin beta subunit A-like domain. The polypeptide, preferably the
XX CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
XX CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
XX CC V at residue 315 and A at residue 320 have been replaced by C, is useful
XX CC for determining if a test compound is a candidate compound for binding to
XX CC CD11b or for treating an inflammatory disorder, by contacting a test
XX CC compound with the polypeptide and determining if the test compound binds
XX CC to the polypeptide. The integrin subunits are useful for reducing
XX CC skeletal muscle injury, for treating disorders caused by ischaemia-
XX CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
XX CC to purify variant integrin polypeptide ligands and as bait proteins in
XX CC two-hybrid or three-hybrid assays. This sequence represents a human
XX CC integrin alpha subunit Alpha 10 variant A domain. Note: This variant
XX CC sequence is not featured in the specification but has been derived from
XX CC the wild-type protein shown in AAU76853
XX XX Sequence 195 AA;
XX Query Match 1.1%; Score 13; DB 5; Length 195;
XX Best Local Similarity 100.0%; Pred. No. 0.00085;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      166 VIVLDGNSIYPW 178
Db      8 VIVLDGNSIYPW 20
|||||
RESULT 71
AAU76853
ID AAU76853 standard; protein; 195 AA.
XX AC AAU76853;
XX DT 21-MAY-2002 (first entry)
XX DE Human integrin alpha subunit Alpha 10 A domain.
XX KW Human; integrin alpha subunit; A domain; Alpha 10; integrin beta subunit;
XX KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
XX KW ischaemia-reperfusion injury; immune complex; parasitic disease;
XX KW antiinflammatory; vasotropic; antiparasitic; vulnery; gene therapy.
XX OS Homo sapiens.
XX XX WO200209737-A1.
XX PN 07-FEB-2002.
XX PF 31-JUL-2001; 2001WO-US023957.
XX PR 31-JUL-2000; 2000US-0221950P.
XX PR 11-JAN-2001; 2001US-00758493.
XX PR 13-MAR-2001; 2001US-00805354.
XX PA (GHEO ) GEN HOSPITAL CORP.
XX PI Arnaout AM, Li R, Xiong J;
XX DR WPI; 2002-188687/24.
XX PT Novel high affinity integrin polypeptide useful for treating restenosis
XX PT and parasitic diseases, comprises all or part of variant integrin alpha
XX PT subunit A domain or variant integrin beta subunit A-like domain.
XX PS Example 2; Fig 5; 55pp; English.
XX CC The invention relates to a high affinity integrin polypeptide comprising
XX CC all or part of a variant integrin alpha subunit A domain or a variant
XX CC integrin beta subunit A-like domain. The polypeptide, preferably the
XX CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
XX CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
XX CC V at residue 315 and A at residue 320 have been replaced by C, is useful
XX CC for determining if a test compound is a candidate compound for binding to
XX CC CD11b or for treating an inflammatory disorder, by contacting a test
XX CC compound with the polypeptide and determining if the test compound binds
XX CC to the polypeptide. The integrin subunits are useful for reducing
XX CC skeletal muscle injury, for treating disorders caused by ischaemia-
XX CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
XX CC to purify variant integrin polypeptide ligands and as bait proteins in
XX CC two-hybrid or three-hybrid assays. This sequence represents the human
XX CC integrin alpha subunit Alpha 10 A domain
XX XX Sequence 195 AA;
XX Query Match 1.1%; Score 13; DB 5; Length 195;
XX Best Local Similarity 100.0%; Pred. No. 0.00085;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 72

AAV32243  
ID AAV32243 standard; protein; 1132 AA.  
XX AC AAY32243;  
XX XX  
DT DT 15-FEB-2000 (first entry)  
DE DE Human integrin subunit alpha-10 splice variant.  
XX XX  
KW Integrin alpha-10; ISal0; human; trauma; rheumatoid arthritis;  
KW osteoarthritis; osteoarthrosis; cancer; atherosclerosis; inflammation;  
KW therapy; cartilage; chondrocyte; osteoblast; fibroblast; vaccine; marker;  
KW splice variant.  
XX XX  
OS Homo sapiens.

|       | Key   | Location/Qualifiers |
|-------|---|---------------------|
| FF FT | Peptide   | 1..22               |
| FT    | /note= "signal peptide"   |                     |
| FT PT | Protein   | 23..1132            |
| FT    | /note= "mature protein"   |                     |
| XX XX |   |                     |
| PN    | WO99S1639-A1.   |                     |
| XX XX |   |                     |
| PD PD | 14-OCT-1999.  |                     |
| XX XX |   |                     |
| PF PF | 31-MAR-1999;  | 99WO-SEROC544.      |
| XX XX |   |                     |
| PR PR | 02-APR-1998;  | 98SE-00001164.      |
| PR PR | 28-JAN-1999;  | 98SE-00000319.      |
| XX XX |   |                     |
| PA PA | (ACTI-) ACTIVE BIOTECH AB.  |                     |
| XX XX |   |                     |
| PI PI | Lundgren-Akerlund E;  |                     |
| XX XX |   |                     |
| DR DR | WPI: 2000-052639/04.  |                     |
| XX XX | N-PSDB; AA234720.   |                     |
| PT PT | New isolated integrin subunit alpha-10, used as a marker or target molecule for cells during development, regeneration and pathological conditions, e.g. arthritis, osteoarthritis, cancer, atherosclerosis or inflammation.  |                     |
| PS PS | Claim 1; Page 43-48; 90pp; English.   |                     |
| XX CC | This sequence represents a splice variant of novel human chondrocyte integrin subunit alpha-10 (ISal0). It is identical to ISal0 (see AAY32242) except for deletion of amino acids 975-986. The invention relates to a recombinant or isolated integrin heterodimer comprising the alpha10 subunit in association with subunit beta (especially beta-1). The heterodimer, subunit alpha-10 or splice variant can be used as a marker or target of all types of cells, e.g. of chondrocytes, osteoblasts and fibroblasts. They can also be used for treating pathological conditions involving ISal0, such as damage to cartilage, trauma, rheumatoid arthritis or osteoarthritis; for detecting the formation of cartilage during embryonal development, physiological or therapeutic repair of cartilage, or detecting regeneration of cartilage or chondrocytes during transplantation of cartilage or chondrocytes; for selection and analysis or for sorting, isolating or purification of chondrocytes and for in vitro studies of differentiation of chondrocytes; and as a target for anti-adhesive drugs or molecules in tendon, ligament, skeletal muscle or other tissues where adhesion impairs the function of the tissue (all claimed). ISal0 binding entities can be used to determine the differentiation-state of cells during embryonic development, angiogenesis or development of cancer, in pathological conditions such as rheumatoid arthritis, osteoarthritis or cancer, in tissue regeneration or in therapeutic and physiological repair of cartilage (claimed). A vaccine comprising the integrin heterodimer or subunit alpha-10 is also claimed. ISal0 polynucleotides, vectors, host cells and methods of producing recombinant ISal0 are also claimed |                     |
| XX SQ | Sequence 1132 AA;   |                     |

RESULT 74  
 AAB64658  
 ID AAB64658 standard; protein; 1152 AA.  
 XX  
 AC AAB64658;  
 XX  
 DT 22-MAR-2001 (first entry)  
 XX  
 DE Human secreted protein BLAST search protein SEQ ID NO: 168.  
 XX  
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;  
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200077197-A1.  
 XX  
 PD 21-DEC-2000.  
 XX  
 PF 01-JUN-2000; 2000WO-US014934.  
 XX  
 PR 11-JUN-1999; 99US-0138599P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.  
 XX  
 PI Rosen CA, Ruben SM, Komatsculis GA;  
 XX  
 DR WPI; 2001-032312/04.  
 XX  
 PT Isolated nucleic acid molecule encoding a human secreted protein is used  
 PT in preventing, treating or ameliorating a medical condition.  
 XX  
 PS Disclosure; Page 547-551; 558pp; English.  
 XX  
 CC The invention relates to the isolation of genes AAF32757-F32803 encoding  
 CC the human secreted proteins AAB64549-B64594. The sequence is used as a  
 CC query sequence for doing BLASTX searches to identify homologous  
 CC sequences. The genes and proteins are useful for preventing, ameliorating  
 CC or treating medical conditions, e.g. by protein or gene therapy. The  
 CC genes are isolated from a range of human tissues disclosed in the  
 CC specification. The nucleic acids, proteins, antibodies and (ant)agonists  
 CC are useful in the diagnosis, treatment and prevention of: (a) cancer,  
 CC e.g. breast and ovarian cancer, and other cancers of the adrenal gland,  
 CC bone, bone marrow, breast, gastrointestinal tract, liver, lung, or  
 CC urogenital; (b) immune disorders e.g. Addison's disease, allergies,  
 CC autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,  
 CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d)  
 CC wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections  
 XX  
 SQ Sequence 1152 AA;  
 Query Match 1.1%; Score 13; DB 4; Length 1152;  
 Best Local Similarity 100.0%; Pred. No. 0.004;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 166 VIVLDGNSIYPW 178  
 DB 169 VIVLDGNSIYPW 181  
 |||||  
 RESULT 75  
 AAY32242  
 ID AAY32242 standard; protein; 1167 AA.  
 XX

AC AAY32242;  
 XX  
 DT 15-FEB-2000 (first entry)  
 XX  
 DE Human integrin subunit alpha-10.  
 XX  
 KW Integrin alpha-10; ISa10; human; trauma; rheumatoid arthritis;  
 KW osteoarthritis; osteoarthritis; cancer; atherosclerosis; inflammation;  
 KW therapy; cartilage; chondrocyte; osteoblast; fibroblast; vaccine; marker.  
 XX  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT Peptide 1..22  
 FT Protein 23..1145 /note= "signal peptide"  
 FT Domain 23..1120 /note= "mature protein"  
 FT Modified-site 98 /note= "extracellular domain"  
 FT Domain 162..359 /note= "N-glycosylated"  
 FT Modified-site 336 /note= "I-domain"  
 FT Modified-site 364 /note= "N-glycosylated"  
 FT Binding-site 494..502 /note= "N-glycosylated"  
 FT Binding-site 558..566 /note= "cation binding site motif"  
 FT Binding-site 620..628 /note= "cation binding site motif"  
 FT Modified-site 733 /note= "cation binding site motif"  
 FT Modified-site 839 /note= "N-glycosylated"  
 FT Modified-site 921 /note= "N-glycosylated"  
 FT Modified-site 1018 /note= "N-glycosylated"  
 FT Modified-site 1039 /note= "N-glycosylated"  
 FT Domain 1121..1145 /note= "transmembrane domain"  
 FT Domain 1122..1167 /note= "cytoplasmic domain, specifically claimed in Claim 21"  
 XX  
 PN WO9951639-A1.  
 XX  
 PD 14-OCT-1999.  
 XX  
 PP 31-MAR-1999; 99WO-SE000544.  
 XX  
 PR 02-APR-1998; 98SE-00001164.  
 PR 28-JAN-1999; 99SE-00000319.  
 XX  
 PA (ACTI-) ACTIVE BIOTECH AB.  
 XX  
 PI Lundgren-Akerlund E;  
 DR WPI; 2000-052639/04.  
 DR N-PSDB; RAZ34719.  
 XX  
 PT New isolated integrin subunit alpha-10, used as a marker or target  
 PT molecule for cells during development, regeneration and pathological  
 PT conditions, e.g. arthritis, osteoarthritis, cancer, atherosclerosis or  
 PT inflammation.  
 XX  
 PS Claim 1; Fig 6; 90pp; English.  
 XX  
 CC This sequence represents novel human chondrocyte integrin subunit alpha-



CC 10 (ISa10). A splice variant is given in AAY32243. The invention relates  
CC to a recombinant or isolated integrin heterodimer comprising the alpha10  
CC subunit in association with subunit beta (especially beta-1). The  
CC heterodimer and the subunit alpha-10 can be used as markers or targets of  
CC all types of cells, e.g. of chondrocytes, osteoblasts and fibroblasts.  
CC They can also be used: for treating pathological conditions involving  
CC ISa10, such as damage to cartilage, trauma, rheumatoid arthritis or  
CC osteoarthritis; for detecting the formation of cartilage during embryonal  
CC development, physiological or therapeutic repair of cartilage, or  
CC detecting regeneration of cartilage or chondrocytes during  
CC transplantation of cartilage or chondrocytes; for selection and analysis  
CC or for sorting, isolating or purification of chondrocytes and for in  
CC vitro studies of differentiation of chondrocytes; and as a target for  
CC anti-adhesive drugs or molecules in tendon, ligament, skeletal muscle or  
CC other tissues where adhesion impairs the function of the tissue (all  
CC claimed). ISa10 binding entities can be used to determine the  
CC differentiation-state of cells during embryonic development, angiogenesis  
CC or development of cancer, in pathological conditions such as rheumatoid  
CC arthritis, osteoarthritis or cancer, in tissue regeneration or in  
CC therapeutic and physiological repair of cartilage (claimed). A  
CC vaccine comprising the integrin heterodimer or subunit alpha-10 is also  
CC claimed. ISa10 polynucleotides, vectors, host cells and methods of  
CC producing recombinant ISa10 are also claimed  
xx  
SQ Sequence 1167 AA;

Query Match 1.18; Score 13; DB 3; Length 1167;  
Best Local Similarity 100.0%; Pred. NO. 0.0041;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 166 VIVLDGNSIYPW 178  
Db 169 VIVLDGNSIYPW 181  
|||||

Search completed: June 24, 2004, 18:03:55  
Job time : 72 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 18:05:07 ; Search time 53 Seconds  
(without alignments)  
6328.087 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 1188

Sequence: 1 MDLPRLVVALSLWPGFT.....FRSARRRREPGLDTPPKVLE 1188

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1163542 seqs, 282313646 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications AA:\*

- 1: /cgm2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgm2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgm2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgm2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgm2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgm2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgm2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgm2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgm2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgm2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgm2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgm2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgm2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgm2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgm2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgm2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgm2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgm2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description         |
|------------|-------|-------------|--------|----|---------------------|
| 1          | 926   | 77.9        | 1188   | 15 | US-10-291-265-338   |
| 2          | 911   | 76.7        | 1188   | 15 | US-10-291-265-810   |
| 3          | 911   | 76.7        | 1189   | 10 | US-09-984-130-35    |
| 4          | 911   | 76.7        | 1189   | 10 | US-09-836-353A-35   |
| 5          | 911   | 76.7        | 1189   | 12 | US-10-262-839-4     |
| 6          | 810   | 68.2        | 1034   | 10 | US-09-984-130-43    |
| 7          | 810   | 68.2        | 1034   | 10 | US-09-836-353A-43   |
| 8          | 753   | 63.4        | 1120   | 12 | US-10-262-839-6     |
| 9          | 489   | 41.2        | 707    | 9  | US-09-764-870-313   |
| 10         | 489   | 41.2        | 707    | 14 | US-10-125-540-313   |
| 11         | 245   | 20.6        | 437    | 15 | US-10-108-260A-3386 |
| 12         | 193   | 16.2        | 193    | 10 | US-09-805-354-8     |
| 13         | 193   | 16.2        | 193    | 12 | US-09-758-493-8     |
| 14         | 193   | 16.2        | 193    | 14 | US-10-144-259-8     |
| 15         | 88    | 7.4         | 103    | 9  | US-09-764-870-472   |

|    |    |     |      |    |                     |                   |
|----|----|-----|------|----|---------------------|-------------------|
| 16 | 88 | 7.4 | 103  | 11 | US-09-764-875-1193  | Sequence 1193, Ap |
| 17 | 88 | 7.4 | 103  | 14 | US-10-125-540-472   | Sequence 472, App |
| 18 | 42 | 3.5 | 42   | 14 | US-10-144-259-25    | Sequence 25, Appl |
| 19 | 42 | 3.5 | 545  | 10 | US-09-866-050A-500  | Sequence 500, App |
| 20 | 42 | 3.5 | 688  | 10 | US-09-866-050A-624  | Sequence 624, App |
| 21 | 42 | 3.5 | 696  | 10 | US-09-866-050A-501  | Sequence 501, App |
| 22 | 33 | 2.8 | 33   | 10 | US-09-984-130-102   | Sequence 102, App |
| 23 | 33 | 2.8 | 33   | 10 | US-09-836-353A-102  | Sequence 102, App |
| 24 | 21 | 1.8 | 158  | 10 | US-09-866-050A-391  | Sequence 391, App |
| 25 | 17 | 1.4 | 85   | 10 | US-09-866-050A-390  | Sequence 390, App |
| 26 | 15 | 1.3 | 148  | 9  | US-09-764-870-284   | Sequence 284, App |
| 27 | 15 | 1.3 | 148  | 9  | US-09-764-870-444   | Sequence 444, App |
| 28 | 15 | 1.3 | 148  | 14 | US-10-125-540-284   | Sequence 284, App |
| 29 | 15 | 1.3 | 148  | 14 | US-10-125-540-444   | Sequence 444, App |
| 30 | 15 | 1.3 | 192  | 12 | US-10-346-863-57    | Sequence 57, Appl |
| 31 | 15 | 1.3 | 192  | 16 | US-10-474-832-59    | Sequence 59, Appl |
| 32 | 15 | 1.3 | 192  | 16 | US-10-474-832-60    | Sequence 60, Appl |
| 33 | 15 | 1.3 | 192  | 16 | US-10-474-832-61    | Sequence 61, Appl |
| 34 | 15 | 1.3 | 195  | 10 | US-09-805-354-5     | Sequence 5, Appl  |
| 35 | 15 | 1.3 | 195  | 12 | US-09-758-493-5     | Sequence 5, Appl  |
| 36 | 15 | 1.3 | 195  | 14 | US-10-144-259-5     | Sequence 5, Appl  |
| 37 | 15 | 1.3 | 212  | 9  | US-09-996-738-5     | Sequence 5, Appl  |
| 38 | 15 | 1.3 | 212  | 9  | US-09-996-738-6     | Sequence 6, Appl  |
| 39 | 15 | 1.3 | 214  | 12 | US-10-625-260-5     | Sequence 5, Appl  |
| 40 | 15 | 1.3 | 214  | 12 | US-10-625-260-6     | Sequence 6, Appl  |
| 41 | 15 | 1.3 | 214  | 12 | US-10-825-260-9     | Sequence 9, Appl  |
| 42 | 15 | 1.3 | 214  | 13 | US-10-061-658-5     | Sequence 5, Appl  |
| 43 | 15 | 1.3 | 214  | 13 | US-10-061-658-6     | Sequence 6, Appl  |
| 44 | 15 | 1.3 | 214  | 13 | US-10-061-658-9     | Sequence 9, Appl  |
| 45 | 15 | 1.3 | 214  | 16 | US-10-474-832-63    | Sequence 63, Appl |
| 46 | 15 | 1.3 | 214  | 16 | US-10-474-832-64    | Sequence 64, Appl |
| 47 | 15 | 1.3 | 1151 | 10 | US-09-984-130-103   | Sequence 103, App |
| 48 | 15 | 1.3 | 1151 | 10 | US-09-836-353A-103  | Sequence 103, App |
| 49 | 15 | 1.3 | 1179 | 12 | US-09-918-715-250   | Sequence 250, App |
| 50 | 15 | 1.3 | 1180 | 12 | US-09-918-715-307   | Sequence 307, App |
| 51 | 13 | 1.1 | 195  | 10 | US-09-805-354-7     | Sequence 7, Appl  |
| 52 | 13 | 1.1 | 195  | 12 | US-09-758-493-7     | Sequence 7, Appl  |
| 53 | 13 | 1.1 | 195  | 14 | US-10-144-259-7     | Sequence 7, Appl  |
| 54 | 11 | 0.9 | 43   | 14 | US-10-144-259-26    | Sequence 26, Appl |
| 55 | 11 | 0.9 | 1147 | 12 | US-10-336-603A-42   | Sequence 42, Appl |
| 56 | 11 | 0.9 | 1181 | 12 | US-10-211-462-187   | Sequence 187, App |
| 57 | 11 | 0.9 | 1181 | 14 | US-10-160-354-2     | Sequence 2, Appl  |
| 58 | 11 | 0.9 | 1181 | 15 | US-10-295-027-1286  | Sequence 1286, Ap |
| 59 | 9  | 0.8 | 164  | 15 | US-10-108-260A-3415 | Sequence 3415, Ap |
| 60 | 9  | 0.8 | 312  | 16 | US-10-038-854-102   | Sequence 102, App |
| 61 | 9  | 0.8 | 371  | 15 | US-10-369-493-2470  | Sequence 2470, Ap |
| 62 | 8  | 0.7 | 17   | 9  | US-09-350-259-20    | Sequence 20, Appl |
| 63 | 8  | 0.7 | 17   | 10 | US-09-891-943-20    | Sequence 20, Appl |
| 64 | 8  | 0.7 | 43   | 14 | US-10-144-259-27    | Sequence 27, Appl |
| 65 | 8  | 0.7 | 67   | 9  | US-09-978-295A-85   | Sequence 85, Appl |
| 66 | 8  | 0.7 | 67   | 9  | US-09-978-697-85    | Sequence 85, Appl |
| 67 | 8  | 0.7 | 67   | 9  | US-09-978-192A-85   | Sequence 85, Appl |
| 68 | 8  | 0.7 | 67   | 9  | US-09-999-832A-85   | Sequence 85, Appl |
| 69 | 8  | 0.7 | 67   | 10 | US-09-378-189-85    | Sequence 85, Appl |
| 70 | 8  | 0.7 | 67   | 10 | US-09-978-608A-85   | Sequence 85, Appl |
| 71 | 8  | 0.7 | 67   | 10 | US-09-978-585A-85   | Sequence 85, Appl |
| 72 | 8  | 0.7 | 67   | 10 | US-09-978-191A-85   | Sequence 85, Appl |
| 73 | 8  | 0.7 | 67   | 10 | US-09-978-403A-85   | Sequence 85, Appl |
| 74 | 8  | 0.7 | 67   | 10 | US-09-978-564A-85   | Sequence 85, Appl |
| 75 | 8  | 0.7 | 67   | 10 | US-09-999-831A-85   | Sequence 85, Appl |
| 76 | 8  | 0.7 | 67   | 10 | US-09-981-915A-85   | Sequence 85, Appl |
| 77 | 8  | 0.7 | 67   | 10 | US-09-978-824-85    | Sequence 85, Appl |
| 78 | 8  | 0.7 | 67   | 10 | US-09-918-585A-85   | Sequence 85, Appl |
| 79 | 8  | 0.7 | 67   | 10 | US-09-978-423A-85   | Sequence 85, Appl |
| 80 | 8  | 0.7 | 67   | 10 | US-09-978-193A-85   | Sequence 85, Appl |
| 81 | 8  | 0.7 | 67   | 10 | US-09-999-830A-85   | Sequence 85, Appl |
| 82 | 8  | 0.7 | 67   | 10 | US-09-978-757A-85   | Sequence 85, Appl |
| 83 | 8  | 0.7 | 67   | 10 | US-09-978-187B-85   | Sequence 85, Appl |
| 84 | 8  | 0.7 | 67   | 10 | US-09-978-643A-85   | Sequence 85, Appl |
| 85 | 8  | 0.7 | 67   | 10 | US-09-978-375A-85   | Sequence 85, Appl |
| 86 | 8  | 0.7 | 67   | 10 | US-09-978-298A-85   | Sequence 85, Appl |
| 87 | 8  | 0.7 | 67   | 10 | US-09-978-188A-85   | Sequence 85, Appl |
| 88 | 8  | 0.7 | 67   | 10 | US-09-978-681A-85   | Sequence 85, Appl |

```
89 8 0.7 67 10 US-09-978-194A-85 Sequence 85, Appl
90 8 0.7 67 10 US-09-999-829A-85 Sequence 85, Appl
91 8 0.7 67 10 US-09-978-299A-85 Sequence 85, Appl
92 8 0.7 67 10 US-09-978-544A-85 Sequence 85, Appl
93 8 0.7 67 10 US-09-978-665A-85 Sequence 85, Appl
94 8 0.7 67 10 US-09-978-802A-85 Sequence 85, Appl
95 8 0.7 67 12 US-10-164-749A-85 Sequence 324, App
96 8 0.7 67 12 US-10-221-278-324 Sequence 85, Appl
97 8 0.7 67 12 US-09-999-831A-85 Sequence 85, Appl
98 8 0.7 67 12 US-10-013-917A-85 Sequence 85, Appl
99 8 0.7 67 12 US-09-999-834A-85 Sequence 85, Appl
100 8 0.7 67 12 US-10-162-521A-85 Sequence 85, Appl

ALIGNMENTS

RESULT 1
US-10-291-265-338
; Sequence 338, Application US/10291265
; Publication No. US20030232054A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides
; CURRENT APPLICATION NUMBER: US/10/291,265
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 338
; LENGTH: 1188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-265-338

Query Match 77.9%; Score 926; DB 15; Length 1188;
Best Local Similarity 99.88; Pred. No. 0;
Matches 1126; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDLPRGLVAVWALS L P G F T D T F N M O T R K P R V I P G S R T A P F G T V Q O H D I S G N K W L V V G A 60
Db 1 MDLPRGLVAVWALS L P G F T D T F N M O T R K P R V I P G S R T A P F G T V Q O H D I S G N K W L V V G A 60

QY 61 PLENTGYQKTDVYKCPVHIGNTKLNGLRVTLISNYSERKDNMRGLGLSLATNPKNPSLA 120
Db 61 PLENTGYQKTDVYKCPVHIGNTKLNGLRVTLISNYSERKDNMRGLGLSLATNPKNPSLA 120

QY 121 CSPLSWHECGSSYTTGMC SRVSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNIYPWVE 180
Db 121 CSPLSWHECGSSYTTGMC SRVSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNIYPWVE 180

QY 181 VQHLINILKRFYIGPQIQGVVQYGEDVVFHSHNDVSVKDVVEAASHIBQRGSTET 240
Db 181 VQHLINILKRFYIGPQIQGVVQYGEDVVFHSHNDVSVKDVVEAASHIBQRGSTET 240

QY 241 RTAFGIEFARSEAFQGRKGAKKMIIVI TDGSHSDSPDLERKVIQOSERDNVTRYAVVL 300
Db 241 RTAFGIEFARSEAFQGRKGAKKMIIVI TDGSHSDSPDLERKVIQOSERDNVTRYAVVL 300

QY 301 GYNNRGINPETELNEIKYIASOPDDKHFNVTDEAALKDIDVALGRIEFLSGTNNKT 360
Db 301 GYNNRGINPETELNEIKYIASOPDDKHFNVTDEAALKDIDVALGRIEFLSGTNNKT 360

QY 361 SFGLENSQTGFSFHHVEDGVLLGAVGAYDNGAVLTKETSAGKVIPLRESYLKEFPEELKN 420
```

```
Db 361 SFGLENSQTGFSFHHVEDGVLLGAVGAYDNGAVLTKETSAGKVIPLRESYLKEFPEELKN 420
QY 421 HGAYLGYYTTSVVSSEQRGVYVAGAPRNHTGKVLFTMHNNESITIHQAMEGQOIGSYF 480
Db 421 HGAYLGYYTTSVVSSEQRGVYVAGAPRNHTGKVLFTMHNNRSITIHQAMEGQOIGSYF 480
QY 481 GSEITSVDIDGSGVTDVLLVGAPMYFNEGRGKGVYVYELRQNRVYVYATGLKDSHSYQVA 540
Db 481 GSEITSVDIDGSGVTDVLLVGAPMYFNEGRGKGVYVYELRQNRVYVYATGLKDSHSYQVA 540
QY 541 RFGSSIASVRDLNQSNDVYVVGAPLEDNHAGAIYIFGFRGSILKTIKQRTASELATG 600
Db 541 RFGSSIASVRDLNQSNDVYVVGAPLEDNHAGAIYIFGFRGSILKTIKQRTASELATG 600
QY 601 LQYFCSSIHGQLDNLNEDGLDLAVGALGNVATLWSRPVVQINASHLHFPSPKINI FHRDCK 660
Db 601 LQYFCSSIHGQLDNLNEDGLDLAVGALGNVATLWSRPVVQINASHLHFPSPKINI FHRDCK 660
QY 661 RSGRDATCLAAFLCPTPIFLAPHFQTTTGVIRYNATDERRYTTPRAHLDGGDRFTNRAV 720
Db 661 RSGRDATCLAAFLCPTPIFLAPHFQTTTGVIRYNATDERRYTTPRAHLDGGDRFTNRAV 720
QY 721 LSSSQBELCERINFHVLTADYVKVPTFSVEYSLEDDPHGPMDDGWPTTLRVSVYFVWG 780
Db 721 LSSSQBELCERINFHVLTADYVKVPTFSVEYSLEDDPHGPMDDGWPTTLRVSVYFVWG 780
QY 781 CNEDEHCVDPDLVDARSOLPTAMEYCORVLRKPAQDCSAYTLSFTTTFVFIIESTRQVAV 840
Db 781 CNEDEHCVDPDLVDARSOLPTAMEYCORVLRKPAQDCSAYTLSFTTTFVFIIESTRQVAV 840
QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSEICWNEERLQKQVCNYSYPP 900
Db 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSEICWNEERLQKQVCNYSYPP 900
QY 901 RAKAKVAFRLDSEFSKSIPLHHLIELAAGSDSNRSTKEDNVAPLPHLYEADVLFT 960
Db 901 RAKAKVAFRLDSEFSKSIPLHHLIELAAGSDSNRSTKEDNVAPLPHLYEADVLFT 960
QY 961 RSSLSHYEVKLNSSLERYDGI GPPSCIFRIONLGLPFIHGMWKITIPITATRSNRL 1020
Db 961 RSSLSHYEVKLNSSLERYDGI GPPSCIFRIONLGLPFIHGMWKITIPITATRSNRL 1020
QY 1021 KLRFELTDEANTSCNMGNSTEYRPTVPEEDLRAPQLNHSNSDVVSINCNIRLVPNQEI 1080
Db 1021 KLRFELTDEANTSCNMGNSTEYRPTVPEEDLRAPQLNHSNSDVVSINCNIRLVPNQEI 1080
QY 1081 NFHLGNLWLSLKALKYKSKMIMVNAALQRFHSPPIFREDDPSROI 1128
Db 1081 NFHLGNLWLSLKALKYKSKMIMVNAALQRFHSPPIFREDDPSROI 1128

RESULT 2
US-10-291-265-810
; Sequence 810, Application US/10291265
; Publication No. US20030232054A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang et al
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-017 (785)
; CURRENT APPLICATION NUMBER: US/10/291,265
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
```

; SEQ ID NO 810  
; LENGTH: 1188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-291-265-810

```
Query Match      76.7%; Score 911; DB 15; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDLPRGLVAVWALSINPFTDFTFNDTRKPRVPGSRTPAFGTYTQQHDSGNKWLAVGA 60
Db 1 MDLPRGLVAVWALSINPFTDFTFNDTRKPRVPGSRTPAFGTYTQQHDSGNKWLAVGA 60
Qy 61 PLETHGYOKTGDVYKCPVHGNCTKLNLRVTLSNVSRKDNMLGLSLATNPKNDSFLA 120
Db 61 PLETHGYOKTGDVYKCPVHGNCTKLNLRVTLSNVSRKDNMLGLSLATNPKNDSFLA 120
Qy 121 CSPLASHCEGSSYYTTCMCSSRVNSFRFSKTVPALQRCQTYMDIVIVLDCSNSIYPWVE 180
Db 121 CSPLASHCEGSSYYTTCMCSSRVNSFRFSKTVPALQRCQTYMDIVIVLDCSNSIYPWVE 180
Qy 181 VQHEFLINILKKFYIGPGQIQGVVQYGEDVWVHEPHLNDYRSVKDVVEAASHIEBORGTTET 240
Db 181 VQHEFLINILKKFYIGPGQIQGVVQYGEDVWVHEPHLNDYRSVKDVVEAASHIEBORGTTET 240
Qy 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGSNDSPLDKVIOQSRDNVTRYAVAVL 300
Db 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGSNDSPLDKVIOQSRDNVTRYAVAVL 300
Qy 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNFVNTDEAALKDIVDALGDRIFSLEGTNNKET 360
Db 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNFVNTDEAALKDIVDALGDRIFSLEGTNNKET 360
Qy 361 SPGLEMSQTGSSHHVVEGVLLGAVGDMNGAVLKTSAKVIFLRSYLKEPPEELKN 420
Db 361 SPGLEMSQTGSSHHVVEGVLLGAVGDMNGAVLKTSAKVIFLRSYLKEPPEELKN 420
Qy 421 HGAYLGTYTTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSITIHQMRGQOIGSYF 480
Db 421 HGAYLGTYTTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSITIHQMRGQOIGSYF 480
Qy 481 RFGSSIASVRDLNQDSYNDVVGAPLENDHAGAIYIFHGFRGSILKTPKQRIITASELATG 600
Db 481 RFGSSIASVRDLNQDSYNDVVGAPLENDHAGAIYIFHGFRGSILKTPKQRIITASELATG 600
Qy 601 LQYFCCSIHQGLDNLNEDGLIDLAVGALGNVILMSRPVVQINASLHFPSPKINIFHRDCK 660
Db 601 LQYFCCSIHQGLDNLNEDGLIDLAVGALGNVILMSRPVVQINASLHFPSPKINIFHRDCK 660
Qy 661 RSGRATCLAAFLCETPIFLAHPHOTTGVIRYNATMDERRYTPRAHLDEGGDRPTNAV 720
Db 661 RSGRATCLAAFLCETPIFLAHPHOTTGVIRYNATMDERRYTPRAHLDEGGDRPTNAV 720
Qy 721 LLSQSQELCERINFHVLDTADYVVKPVTFSVYSLEDPDHGPMLDDGWPMTTLRVSVPFNG 780
Db 721 LLSQSQELCERINFHVLDTADYVVKPVTFSVYSLEDPDHGPMLDDGWPMTTLRVSVPFNG 780
Qy 781 CNEDEHCVDPDLVDARSPLPTAMEYCORVLRKPAQDCSAYTSLRDTTVPFIITESTRQYAV 840
Db 781 CNEDEHCVDPDLVDARSPLPTAMEYCORVLRKPAQDCSAYTSLRDTTVPFIITESTRQYAV 840
Qy 841 EATLENRGENAYSTVLNLSQSANLQFASLIQKEDSDGSEIECNBERLQKQCNVSPFFP 900
Db 841 EATLENRGENAYSTVLNLSQSANLQFASLIQKEDSDGSEIECNBERLQKQCNVSPFFP 900
Qy 901 RAKAKVAFRLD 911
Db 901 RAKAKVAFRLD 911
```

RESULT 3  
US-09-984-130-35  
; Sequence 35, Application US/09984130  
; Publication No. US20030055231A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 12 Human Secreted Proteins  
; FILE REFERENCE: PF489P2  
; CURRENT APPLICATION NUMBER: US/09/984.130  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/243,792  
; PRIOR FILING DATE: 2000-10-30  
; PRIOR APPLICATION NUMBER: 09/836,353  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/198,407  
; PRIOR FILING DATE: 2000-04-19  
; PRIOR APPLICATION NUMBER: PCT/US99/25031  
; PRIOR FILING DATE: 1999-10-27  
; PRIOR APPLICATION NUMBER: 60/105,971  
; PRIOR FILING DATE: 1998-10-28  
; NUMBER OF SEQ ID NOS: 149  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 35  
; LENGTH: 1189  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-984-130-35

```
Query Match      76.7%; Score 911; DB 10; Length 1189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDLPRGLVAVWALSINPFTDFTFNDTRKPRVPGSRTPAFGTYTQQHDSGNKWLAVGA 60
Db 1 MDLPRGLVAVWALSINPFTDFTFNDTRKPRVPGSRTPAFGTYTQQHDSGNKWLAVGA 60
Qy 61 PLETHGYOKTGDVYKCPVHGNCTKLNLRVTLSNVSRKDNMLGLSLATNPKNDSFLA 120
Db 61 PLETHGYOKTGDVYKCPVHGNCTKLNLRVTLSNVSRKDNMLGLSLATNPKNDSFLA 120
Qy 121 CSPLASHCEGSSYYTTCMCSSRVNSFRFSKTVPALQRCQTYMDIVIVLDCSNSIYPWVE 180
Db 121 CSPLASHCEGSSYYTTCMCSSRVNSFRFSKTVPALQRCQTYMDIVIVLDCSNSIYPWVE 180
Qy 181 VQHEFLINILKKFYIGPGQIQGVVQYGEDVWVHEPHLNDYRSVKDVVEAASHIEBORGTTET 240
Db 181 VQHEFLINILKKFYIGPGQIQGVVQYGEDVWVHEPHLNDYRSVKDVVEAASHIEBORGTTET 240
Qy 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGSNDSPLDKVIOQSRDNVTRYAVAVL 300
Db 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGSNDSPLDKVIOQSRDNVTRYAVAVL 300
Qy 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNFVNTDEAALKDIVDALGDRIFSLEGTNNKET 360
Db 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNFVNTDEAALKDIVDALGDRIFSLEGTNNKET 360
Qy 361 SPGLEMSQTGSSHHVVEGVLLGAVGDMNGAVLKTSAKVIFLRSYLKEPPEELKN 420
Db 361 SPGLEMSQTGSSHHVVEGVLLGAVGDMNGAVLKTSAKVIFLRSYLKEPPEELKN 420
Qy 421 HGAYLGTYTTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSITIHQMRGQOIGSYF 480
Db 421 HGAYLGTYTTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSITIHQMRGQOIGSYF 480
Qy 481 GSEITTSVDIDGQGVTDVLLVGA PMYFNEGRGKVVYVYELRQNRFFVNGTLKDSHSYQNA 540
Db 481 GSEITTSVDIDGQGVTDVLLVGA PMYFNEGRGKVVYVYELRQNRFFVNGTLKDSHSYQNA 540
Qy 541 RFGSSIASVRDLNQDSYNDVVGAPLENDHAGAIYIFHGFRGSILKTPKQRIITASELATG 600
Db 541 RFGSSIASVRDLNQDSYNDVVGAPLENDHAGAIYIFHGFRGSILKTPKQRIITASELATG 600
```

```

QY 601 LQYFGCSIHGOLDLNEGLDILAVGALGNVILWSRPVQVQINASLHFEPSKINIPHRDCK 660
DB 601 LQYFGCSIHGOLDLNEGLDILAVGALGNVILWSRPVQVQINASLHFEPSKINIPHRDCK 660
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDGGDRFTNRVAV 720
DB 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDGGDRFTNRVAV 720
QY 721 LSSGQELCERINHFVLDADYVYKPTVTSVEYSLEDPDHG2MLDDGMPPTLVSVPFWNG 780
DB 721 LSSGQELCERINHFVLDADYVYKPTVTSVEYSLEDPDHG2MLDDGMPPTLVSVPFWNG 780
QY 781 CNEDEHCVPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTILSFDTTVEIESTRQVAV 840
DB 781 CNEDEHCVPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTILSFDTTVEIESTRQVAV 840
QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECTVNEERLQKQVCNVSYPFF 900
DB 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECTVNEERLQKQVCNVSYPFF 900
QY 901 RAKAKVAFRLD 911
DB 901 RAKAKVAFRLD 911

```

RESULT 4

US-09-836-353A-35  
; Sequence 35, Application US/09836353A  
; Publication No. US20030129685A1

GENERAL INFORMATION:

; APPLICANT: Ni et al.

; TITLE OF INVENTION: 12 Human Secreted Proteins

; FILE REFERENCE: PF489P1

; CURRENT APPLICATION NUMBER: US/09/836,353A

; CURRENT FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/198,407

; PRIOR FILING DATE: 2000-04-19

; PRIOR APPLICATION NUMBER: PCT/US99/25031

; PRIOR FILING DATE: 1999-10-27

; PRIOR FILING DATE: 1998-10-28

; NUMBER OF SEQ ID NOS: 147

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 35

; LENGTH: 1189

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-836-353A-35

Query Match 76.7%; Score 911; DB 10; Length 1189;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MDLPRLGVAVWALSMPGFTDTFNMOTRKPRIIPGRTAFFGYTVOQHDISGNKMLVGA 60
DB 1 MDLPRLGVAVWALSMPGFTDTFNMOTRKPRIIPGRTAFFGYTVOQHDISGNKMLVGA 60
QY 61 PLETNGYOKTGDVYKCPVHNGCTKLNLRVTLSNVSEKDNRLGLSLATNPKNLSPLA 120
DB 61 PLETNGYOKTGDVYKCPVHNGCTKLNLRVTLSNVSEKDNRLGLSLATNPKNLSPLA 120
QY 121 CSPLWSEHCSSYTTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVILDGNSIYPWVE 180
DB 121 CSPLWSEHCSSYTTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVILDGNSIYPWVE 180
QY 181 VOHFLINILKFFVIGGQIQGVQGVQGVQGVQGVQGVQGVQGVQGVQGVQGVQGVQGV 240
DB 181 VOHFLINILKFFVIGGQIQGVQGVQGVQGVQGVQGVQGVQGVQGVQGVQGVQGV 240
QY 241 RTAFGEFARSFAFGKGRGKAKVMIVITDGSHESDSPLEKVIQOSERDNVTRYAVAVL 300
DB 241 RTAFGEFARSFAFGKGRGKAKVMIVITDGSHESDSPLEKVIQOSERDNVTRYAVAVL 300

```

RESULT 5

US-10-262-839-4

; Sequence 4, Application US/10262839

; Publication No. US20040038877A1

GENERAL INFORMATION:

; APPLICANT: Alsbrook, John.

; APPLICANT: Anderson, David W.,

; APPLICANT: Boldog, Ferenc,

; APPLICANT: Burgess, Catherine,

; APPLICANT: Catterton, Elina,

; APPLICANT: Edinger, Shlomit,

; APPLICANT: Ellerman, Karen,

; APPLICANT: Gerlach, Valerie,

; APPLICANT: Gorman, Linda,

; APPLICANT: Guo, Xiaojia,

; APPLICANT: Ji, Weizhen,

; APPLICANT: Kekuda, Ramesh,

; APPLICANT: Leach, Martin,

; APPLICANT: Li, Li,

; APPLICANT: Miller, Charles,

; APPLICANT: Patturajan, Meera,

; APPLICANT: Reiger, Daniel,

; APPLICANT: Rothenberg, Mark,

; APPLICANT: Shinkets, Richard,

; APPLICANT: Smithson, Glennda,

; APPLICANT: Spyttek, Kimberly,

; APPLICANT: Taupier, Raymond, Jr.,

```

; APPLICANT: Vernet, Corine,
; APPLICANT: Voss, Edward,
; APPLICANT: Zethusen, Brian,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-462A
; CURRENT APPLICATION NUMBER: US/10/262,839
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,101
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/371,972
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/327,342
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/328,044
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/374,738
; PRIOR FILING DATE: 2002-04-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: Curaseq1st version 0.1
; SEQ ID NO 4
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-839-4

```

```

Query Match 76.7%; Score 911; DB 12; Length 1189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLPRLVAVALSLLPQGTDTFNDTRKPRVPGSRPTAFPGYTVQOHDISGNKWLAVGA 60
DB 1 MDLPRLVAVALSLLPQGTDTFNDTRKPRVPGSRPTAFPGYTVQOHDISGNKWLAVGA 60
QY 61 PLETNGYQKTGDVYKCPVHGNCTKLNIGRVTLNVSERKNNRGLSLATNPKNNSFLA 120
DB 61 PLETNGYQKTGDVYKCPVHGNCTKLNIGRVTLNVSERKNNRGLSLATNPKNNSFLA 120
QY 121 CSPLMSHECGSSYYTGMCNRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
DB 121 CSPLMSHECGSSYYTGMCNRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
QY 181 VOHFLNLIKKEYPGQIOGVWQYQGVQEDVVEHPLNDYRSVKQVVEAASHIEQGGTET 240
DB 181 VOHFLNLIKKEYPGQIOGVWQYQGVQEDVVEHPLNDYRSVKQVVEAASHIEQGGTET 240
QY 241 RTAFGIEFARSAFQKGRKGAKKVMIVITDGEHSDSPDLKVIQOQSERDNVTRYAVAVL 300
DB 241 RTAFGIEFARSAFQKGRKGAKKVMIVITDGEHSDSPDLKVIQOQSERDNVTRYAVAVL 300
QY 301 GYNRRGINPETFLNEIKYIASDDDDKGFNNVTDEAALKDVIDALGDRIFPLEGNTNNET 360
DB 301 GYNRRGINPETFLNEIKYIASDDDDKGFNNVTDEAALKDVIDALGDRIFPLEGNTNNET 360
QY 361 SPGLEMSQTGFSSHHVDEGVLLGAVGAYDNGAVLKETSAKVTPLRESYLKBPPEELKN 420
DB 361 SPGLEMSQTGFSSHHVDEGVLLGAVGAYDNGAVLKETSAKVTPLRESYLKBPPEELKN 420
QY 421 HGAYLGTYTWSVSSRQGRVYVAGAPRNFHTGKVLFTPMNNRSLTIHQAVRGOQOIGSYF 480
DB 421 HGAYLGTYTWSVSSRQGRVYVAGAPRNFHTGKVLFTPMNNRSLTIHQAVRGOQOIGSYF 480

```

```

QY 481 GSEITSVDIDGQVTDVLLVAGAPMYFNEGERGKVVYVELAQNRFYVNGTLKDSHSYQNA 540
DB 481 GSEITSVDIDGQVTDVLLVAGAPMYFNEGERGKVVYVELAQNRFYVNGTLKDSHSYQNA 540
QY 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNEAGAIYIFHGFSGSLKTPKQKITASELATG 600
DB 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNEAGAIYIFHGFSGSLKTPKQKITASELATG 600
QY 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRVVQINASLHPEPSKINIFHRDCK 660
DB 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRVVQINASLHPEPSKINIFHRDCK 660
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTPRAHLDEGGDPTNRVAV 720
DB 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTPRAHLDEGGDPTNRVAV 720
QY 721 LLSSGQELCERINRFHVLDATDYVKPVTFSVEYSLEDDPHGFMPLDDGHPFTTLRVSPFWNG 780
DB 721 LLSSGQELCERINRFHVLDATDYVKPVTFSVEYSLEDDPHGFMPLDDGHPFTTLRVSPFWNG 780
QY 781 CNEDEHCVPLDVLDAKSDPLTAMEYCORVLRKPAQDCSAYTLSDTTVFIIESTRQAVAV 840
DB 781 CNEDEHCVPLDVLDAKSDPLTAMEYCORVLRKPAQDCSAYTLSDTTVFIIESTRQAVAV 840
QY 841 EATLENERGENAYSTVLMISQSANLQFASLIQKEDSDGSIKCNWEERLQKQVCNVSYPFF 900
DB 841 EATLENERGENAYSTVLMISQSANLQFASLIQKEDSDGSIKCNWEERLQKQVCNVSYPFF 900
QY 901 RAKAKVAFRLD 911
DB 901 RAKAKVAFRLD 911

```

```

RESULT 6
US-09-984-130-43
; Sequence 43, Application US/09984130
; Publication No. US2003005231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: P2489P2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 43
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-43

```

```

Query Match 68.2%; Score 810; DB 10; Length 1034;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDLPRLVAVALSLLPQGTDTFNDTRKPRVPGSRPTAFPGYTVQOHDISGNKWLAVGA 60
DB 1 MDLPRLVAVALSLLPQGTDTFNDTRKPRVPGSRPTAFPGYTVQOHDISGNKWLAVGA 60
QY 61 PLETNGYQKTGDVYKCPVHGNCTKLNIGRVTLNVSERKNNRGLSLATNPKNNSFLA 120
DB 61 PLETNGYQKTGDVYKCPVHGNCTKLNIGRVTLNVSERKNNRGLSLATNPKNNSFLA 120

```

QY 121 CSPWASHCEGSSYYTTCMCVRNSNFRFSKTVAPALQRCOTYMDIVIVLDGNSNIYPWVE 180  
 DB 121 CSPWASHCEGSSYYTTCMCVRNSNFRFSKTVAPALQRCOTYMDIVIVLDGNSNIYPWVE 180  
 QY 181 VQFLINILKFKFYIGPQIOGVVQVGEDVWVHEFLNDYRSVKOVVRAASHIEQGGTET 240  
 DB 181 VQFLINILKFKFYIGPQIOGVVQVGEDVWVHEFLNDYRSVKOVVRAASHIEQGGTET 240  
 QY 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSHPDPLEKVIQOESRDNVTRYAVAVL 300  
 DB 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSHPDPLEKVIQOESRDNVTRYAVAVL 300  
 QY 301 GYNRRGINPETPNEIKYIASDPDDKHFFNVITDEAALKDIDVALGDRIFSLGKTKNET 360  
 DB 301 GYNRRGINPETPNEIKYIASDPDDKHFFNVITDEAALKDIDVALGDRIFSLGKTKNET 360  
 QY 361 SFGLMSQTFSSSHVVEDGVLLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEPPELKN 420  
 DB 361 SFGLMSQTFSSSHVVEDGVLLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEPPELKN 420  
 QY 421 HGAYLGYYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMNNRSLTIHQMRGQOIGSYF 480  
 DB 421 HGAYLGYYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMNNRSLTIHQMRGQOIGSYF 480  
 QY 481 GSEITSDIDGVTDLVLLGAPMYFNEGRGKVVYVELRQNRVYNGTILKDSHSYQNA 540  
 DB 481 GSEITSDIDGVTDLVLLGAPMYFNEGRGKVVYVELRQNRVYNGTILKDSHSYQNA 540  
 QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFSGSILKTPKORITASELATG 600  
 DB 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFSGSILKTPKORITASELATG 600  
 QY 601 LQYFGCSIHGQDLNEDGLIDVALGAGNAILWSRPVQINASLHPEPSKINIFHRDCK 660  
 DB 601 LQYFGCSIHGQDLNEDGLIDVALGAGNAILWSRPVQINASLHPEPSKINIFHRDCK 660  
 QY 661 RSGRDATCLAAFLCFTPIFLAPHQTTVGIRYNATMDERRYTPRAHLDGEGDRFTNRAV 720  
 DB 661 RSGRDATCLAAFLCFTPIFLAPHQTTVGIRYNATMDERRYTPRAHLDGEGDRFTNRAV 720  
 QY 721 LLSGGQELCERINFHVLDTADYVKPTFSVEYSLEDDPHGPMDDGWTTLRVSVFPWNG 780  
 DB 721 LLSGGQELCERINFHVLDTADYVKPTFSVEYSLEDDPHGPMDDGWTTLRVSVFPWNG 780  
 QY 781 CNEDEHCVDPDLVDARSDLPTAMEYQORVLRKPAQDCSAYTILSFDTTFFIESTRQAV 840  
 DB 781 CNEDEHCVDPDLVDARSDLPTAMEYQORVLRKPAQDCSAYTILSFDTTFFIESTRQAV 840  
 QY 841 EATLENRGENAYSTVLNISQSANLQFASLLOKEDSDGSECTVNEERLQKQVNCVSPFF 900  
 DB 841 EATLENRGENAYSTVLNISQSANLQFASLLOKEDSDGSECTVNEERLQKQVNCVSPFF 900  
 QY 901 RAKAKVAFRLD 911  
 DB 901 RAKAKVAFRLD 911

RESULT 7

US-09-836-353A-43  
 ; Sequence 43, Application US/09836353A  
 ; Publication No US20030129685A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NI et al.  
 ; TITLE OF INVENTION: 12 Human Secreted Proteins  
 ; FILE REFERENCE: PF499P1  
 ; CURRENT APPLICATION NUMBER: US/09/836,353A  
 ; CURRENT FILING DATE: 2001-04-18  
 ; PRIOR APPLICATION NUMBER: 60/198,407  
 ; PRIOR FILING DATE: 2000-04-19  
 ; PRIOR APPLICATION NUMBER: PCT/US99/25031  
 ; PRIOR FILING DATE: 1999-10-27  
 ; PRIOR APPLICATION NUMBER: 60/105,971  
 ; PRIOR FILING DATE: 1998-10-28

; NUMBER OF SEQ ID NOS: 147  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 43  
 ; LENGTH: 1034  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-836-353A-43

Query Match 68.2%; Score 810; DB 10; Length 1034;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDLPRLGVAVWALSILWPGFTDTENMDTKPRVIPCSTAPRGYVVOQHDISGNKRLVWGA 60  
 DB 1 MDLPRLGVAVWALSILWPGFTDTENMDTKPRVIPCSTAPRGYVVOQHDISGNKRLVWGA 60  
 QY 61 PLETTYGYOKTGDVYKCPVIHGNCTKLALGRVTLNSVSEKDNMRGLGSLATNPRDNSFLA 120  
 DB 61 PLETTYGYOKTGDVYKCPVIHGNCTKLALGRVTLNSVSEKDNMRGLGSLATNPRDNSFLA 120  
 QY 121 CSPWASHCEGSSYYTTCMCVRNSNFRFSKTVAPALQRCOTYMDIVIVLDGNSNIYPWVE 180  
 DB 121 CSPWASHCEGSSYYTTCMCVRNSNFRFSKTVAPALQRCOTYMDIVIVLDGNSNIYPWVE 180  
 QY 181 VQFLINILKFKFYIGPQIOGVVQVGEDVWVHEFLNDYRSVKOVVRAASHIEQGGTET 240  
 DB 181 VQFLINILKFKFYIGPQIOGVVQVGEDVWVHEFLNDYRSVKOVVRAASHIEQGGTET 240  
 QY 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSHPDPLEKVIQOESRDNVTRYAVAVL 300  
 DB 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSHPDPLEKVIQOESRDNVTRYAVAVL 300  
 QY 301 GYNRRGINPETPNEIKYIASDPDDKHFFNVITDEAALKDIDVALGDRIFSLGKTKNET 360  
 DB 301 GYNRRGINPETPNEIKYIASDPDDKHFFNVITDEAALKDIDVALGDRIFSLGKTKNET 360  
 QY 361 SFGLMSQTFSSSHVVEDGVLLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEPPELKN 420  
 DB 361 SFGLMSQTFSSSHVVEDGVLLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEPPELKN 420  
 QY 421 HGAYLGYYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMNNRSLTIHQMRGQOIGSYF 480  
 DB 421 HGAYLGYYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMNNRSLTIHQMRGQOIGSYF 480  
 QY 481 GSEITSDIDGVTDLVLLGAPMYFNEGRGKVVYVELRQNRVYNGTILKDSHSYQNA 540  
 DB 481 GSEITSDIDGVTDLVLLGAPMYFNEGRGKVVYVELRQNRVYNGTILKDSHSYQNA 540  
 QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFSGSILKTPKORITASELATG 600  
 DB 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFSGSILKTPKORITASELATG 600  
 QY 601 LQYFGCSIHGQDLNEDGLIDVALGAGNAILWSRPVQINASLHPEPSKINIFHRDCK 660  
 DB 601 LQYFGCSIHGQDLNEDGLIDVALGAGNAILWSRPVQINASLHPEPSKINIFHRDCK 660  
 QY 661 RSGRDATCLAAFLCFTPIFLAPHQTTVGIRYNATMDERRYTPRAHLDGEGDRFTNRAV 720  
 DB 661 RSGRDATCLAAFLCFTPIFLAPHQTTVGIRYNATMDERRYTPRAHLDGEGDRFTNRAV 720  
 QY 721 LLSGGQELCERINFHVLDTADYVKPTFSVEYSLEDDPHGPMDDGWTTLRVSVFPWNG 780  
 DB 721 LLSGGQELCERINFHVLDTADYVKPTFSVEYSLEDDPHGPMDDGWTTLRVSVFPWNG 780  
 QY 781 CNEDEHCVDPDLVDARSDLPTAMEYQORVLRKPAQDCSAYTILSFDTTFFIESTRQAV 840  
 DB 781 CNEDEHCVDPDLVDARSDLPTAMEYQORVLRKPAQDCSAYTILSFDTTFFIESTRQAV 840  
 QY 841 EATLENRGENAYSTVLNISQSANLQFASLLOKEDSDGSECTVNEERLQKQVNCVSPFF 900  
 DB 841 EATLENRGENAYSTVLNISQSANLQFASLLOKEDSDGSECTVNEERLQKQVNCVSPFF 900  
 QY 901 RAKAKVAFRLD 911



Db 901 RAKAKVAFRLD 911

|||||

Db 90 CQTMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVGVQGVVHVEPHLND 149

Qy 219 YRSVKDWEAAASHIEBGGTETRTAFGIEPARSEAFQKGRKGAKKVMIVITGESHDSP 278

Db 150 YRSVKDWEAAASHIEBGGTETRTAFGIEPARSEAFQKGRKGAKKVMIVITGESHDSP 209

Qy 279 DLEKVIQOESRDVTRVAVAVLGYNNRGINPTFLNEIKYIASDDPKHFFNVDTBAAL 338

Db 210 DLEKVIQOESRDVTRVAVAVLGYNNRGINPTFLNEIKYIASDDPKHFFNVDTBAAL 269

Qy 339 KDIVDALGDRIFSLEGTNKNETSFGLEMSQTGSSHVVEDGVLLGAVGADMGAVUKET 398

Db 270 KDIVDALGDRIFSLEGTNKNETSFGLEMSQTGSSHVVEDGVLLGAVGADMGAVUKET 329

Qy 399 SAGKVIPLRESYLKPEPEELKNEGAYLGYTVTSVSSROGRVYVAGAPRHHGKVLFT 458

Db 330 SAGKVIPLRESYLKPEPEELKNEGAYLGYTVTSVSSROGRVYVAGAPRHHGKVLFT 389

Qy 459 MHNRSITIHQAMRGQIGSYFGSEITSDIDGQVTDVLLVGA PMYFNEGRERKGVVY 518

Db 390 MHNRSITIHQAMRGQIGSYFGSEITSDIDGQVTDVLLVGA PMYFNEGRERKGVVY 449

Qy 519 ELRQNRVYNGTILKDSHYQNARPGSSIASVRDINODSYNDVAVGAPLEDNHAGAIYIFH 578

Db 450 ELRQNRVYNGTILKDSHYQNARPGSSIASVRDINODSYNDVAVGAPLEDNHAGAIYIFH 509

Qy 579 GFRGSIKTPKQITASELATGLQYFGCSIHGQDLNEDGLIDLVAGALGNVILWRPV 638

Db 510 GFRGSIKTPKQITASELATGLQYFGCSIHGQDLNEDGLIDLVAGALGNVILWRPV 569

Qy 639 VOINASHPEPSKINIFHRDCKSGRDATCLAAFLCTPIFLAPHFQTTVGRYNAATMD 698

Db 570 VOINASHPEPSKINIFHRDCKSGRDATCLAAFLCTPIFLAPHFQTTVGRYNAATMD 629

Qy 699 ERYTTPRAHLDEGGDRFTNRAVLSSGQELCERINHFVLDADYVKVPTFSVEYSLEDDPD 758

Db 630 ERYTTPRAHLDEGGDRFTNRAVLSSGQELCERINHFVLDADYVKVPTFSVEYSLEDDPD 689

Qy 759 HGPMLDGGMPTTLRVSVFPWNGCNEDEHCVDPDLVLDARSDLPTAMEYQCRVLRKPAQDCS 818

Db 690 HGPMLDGGMPTTLRVSVFPWNGCNEDEHCVDPDLVLDARSDLPTAMEYQCRVLRKPAQDCS 749

Qy 819 AYTLSPDFTTVFIESTRQVRVAVATLENRGENAYSTVNLISQSANLQFASLIQKEDSDGS 878

Db 750 AYTLSPDFTTVFIESTRQVRVAVATLENRGENAYSTVNLISQSANLQFASLIQKEDSDGS 809

Qy 879 IECVNEERLQKQVCNVSVYFFFRAXAKVAFRLD 911

Db 810 IECVNEERLQKQVCNVSVYFFFRAXAKVAFRLD 842

RESULT 9

US-09-764-870-313

; Sequence 313, Application US/09764870

; Patent No. US20020042386A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PTZ14

; CURRENT APPLICATION NUMBER: US/09/764,870

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 646

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 313

; LENGTH: 707

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-764-870-313

Query Match 41.2%; Score 489; DB 9; Length 707;

Best Local Similarity 99.7%; Pred. No. 0;

Db 901 RAKAKVAFRLD 911

|||||

Db 90 CQTMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVGVQGVVHVEPHLND 149

Qy 219 YRSVKDWEAAASHIEBGGTETRTAFGIEPARSEAFQKGRKGAKKVMIVITGESHDSP 278

Db 150 YRSVKDWEAAASHIEBGGTETRTAFGIEPARSEAFQKGRKGAKKVMIVITGESHDSP 209

Qy 279 DLEKVIQOESRDVTRVAVAVLGYNNRGINPTFLNEIKYIASDDPKHFFNVDTBAAL 338

Db 210 DLEKVIQOESRDVTRVAVAVLGYNNRGINPTFLNEIKYIASDDPKHFFNVDTBAAL 269

Qy 339 KDIVDALGDRIFSLEGTNKNETSFGLEMSQTGSSHVVEDGVLLGAVGADMGAVUKET 398

Db 270 KDIVDALGDRIFSLEGTNKNETSFGLEMSQTGSSHVVEDGVLLGAVGADMGAVUKET 329

Qy 399 SAGKVIPLRESYLKPEPEELKNEGAYLGYTVTSVSSROGRVYVAGAPRHHGKVLFT 458

Db 330 SAGKVIPLRESYLKPEPEELKNEGAYLGYTVTSVSSROGRVYVAGAPRHHGKVLFT 389

Qy 459 MHNRSITIHQAMRGQIGSYFGSEITSDIDGQVTDVLLVGA PMYFNEGRERKGVVY 518

Db 390 MHNRSITIHQAMRGQIGSYFGSEITSDIDGQVTDVLLVGA PMYFNEGRERKGVVY 449

Qy 519 ELRQNRVYNGTILKDSHYQNARPGSSIASVRDINODSYNDVAVGAPLEDNHAGAIYIFH 578

Db 450 ELRQNRVYNGTILKDSHYQNARPGSSIASVRDINODSYNDVAVGAPLEDNHAGAIYIFH 509

Qy 579 GFRGSIKTPKQITASELATGLQYFGCSIHGQDLNEDGLIDLVAGALGNVILWRPV 638

Db 510 GFRGSIKTPKQITASELATGLQYFGCSIHGQDLNEDGLIDLVAGALGNVILWRPV 569

Qy 639 VOINASHPEPSKINIFHRDCKSGRDATCLAAFLCTPIFLAPHFQTTVGRYNAATMD 698

Db 570 VOINASHPEPSKINIFHRDCKSGRDATCLAAFLCTPIFLAPHFQTTVGRYNAATMD 629

Qy 699 ERYTTPRAHLDEGGDRFTNRAVLSSGQELCERINHFVLDADYVKVPTFSVEYSLEDDPD 758

Db 630 ERYTTPRAHLDEGGDRFTNRAVLSSGQELCERINHFVLDADYVKVPTFSVEYSLEDDPD 689

Qy 759 HGPMLDGGMPTTLRVSVFPWNGCNEDEHCVDPDLVLDARSDLPTAMEYQCRVLRKPAQDCS 818

Db 690 HGPMLDGGMPTTLRVSVFPWNGCNEDEHCVDPDLVLDARSDLPTAMEYQCRVLRKPAQDCS 749

Qy 819 AYTLSPDFTTVFIESTRQVRVAVATLENRGENAYSTVNLISQSANLQFASLIQKEDSDGS 878

Db 750 AYTLSPDFTTVFIESTRQVRVAVATLENRGENAYSTVNLISQSANLQFASLIQKEDSDGS 809

Qy 879 IECVNEERLQKQVCNVSVYFFFRAXAKVAFRLD 911

Db 810 IECVNEERLQKQVCNVSVYFFFRAXAKVAFRLD 842

FILE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD

FILE REFERENCE: 21402-462A

CURRENT APPLICATION NUMBER: US/10/262,839

CURRENT FILING DATE: 2002-10-01

PRIOR APPLICATION NUMBER: 60/326,483

PRIOR FILING DATE: 2001-10-02

PRIOR APPLICATION NUMBER: 60/327,917

PRIOR FILING DATE: 2001-10-09

PRIOR APPLICATION NUMBER: 60/328,029

PRIOR FILING DATE: 2001-10-09

PRIOR APPLICATION NUMBER: 60/328,056

PRIOR FILING DATE: 2001-10-09

PRIOR APPLICATION NUMBER: 60/381,101

PRIOR FILING DATE: 2002-05-16

PRIOR APPLICATION NUMBER: 60/371,972

PRIOR FILING DATE: 2002-04-12

PRIOR APPLICATION NUMBER: 60/327,342

PRIOR FILING DATE: 2001-10-05

PRIOR APPLICATION NUMBER: 60/328,044

PRIOR FILING DATE: 2001-10-09

PRIOR APPLICATION NUMBER: 60/328,849

PRIOR FILING DATE: 2001-10-12

PRIOR APPLICATION NUMBER: 60/374,738

PRIOR FILING DATE: 2002-04-23

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 367

SOFTWARE: CuraSeqlist version 0.1

SEQ ID NO 6

LENGTH: 1120

TYPE: PRT

ORGANISM: Homo sapiens

US-10-262-839-6

Query Match 63.4%; Score 753; DB 12; Length 1120;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQTMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVGVQGVVHVEPHLND 218

Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MDLPRGLVVAWALSMPGFTDTFNMDTRKPRVPGSRTAFPGYTVQOHDISGNKWLWVGA 60  
 Db 17 MDLPRGLVVAWALSMPGFTDTFNMDTRKPRVPGSRTAFPGYTVQOHDISGNKWLWVGA 76

Qy 61 PLETNQYQKTGDVYKCPVHGNCTKLNLRVTLNVSERKDNMRGLSLATNPKDNSFLA 120  
 Db 77 PLETNQYQKTGDVYKCPVHGNCTKLNLRVTLNVSERKDNMRGLSLATNPKDNSFLA 136

Qy 121 CSPLWSHECGSSYYTTGCMRSRVNSFRSKTVAPALQRCQTYMDIVVLDSGNSIYPWVE 180  
 Db 137 CSPLWSHECGSSYYTTGCMRSRVNSFRSKTVAPALQRCQTYMDIVVLDSGNSIYPWVE 196

Qy 181 VQHFLLNLKXFFIYGPQIQGVVQVGEDVWHEFLNDYRSVKDVVEAASHIEQGGTET 240  
 Db 197 VQHFLLNLKXFFIYGPQIQGVVQVGEDVWHEFLNDYRSVKDVVEAASHIEQGGTET 256

Qy 241 RTAFGIEFARSEAPKQGRGAKKVMIVITDGHSDSPDLKVIQOESRDNVTRYAVAVL 300  
 Db 257 RTAFGIEFARSEAPKQGRGAKKVMIVITDGHSDSPDLKVIQOESRDNVTRYAVAVL 316

Qy 301 GYNNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSLGNTKNET 360  
 Db 317 GYNNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSLGNTKNET 376

Qy 361 SPGLEMSQTGPSSHVDEGVLLGAVGAYDWNGAVLKTSGAKVILPRESYLKEPPEELKN 420  
 Db 377 SPGLEMSQTGPSSHVDEGVLLGAVGAYDWNGAVLKTSGAKVILPRESYLKEPPEELKN 436

Qy 421 HGAYLGYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMHNNRSLTIHQMRGQOIGSYF 480  
 Db 437 HGAYLGYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMHNNRSLTIHQMRGQOIGSYF 496

Qy 481 GSEITSDVDGDTVLLVAGAPMYFNEGRERKGVYVYELRQNRVYNGTLKDSHYQNA 540  
 Db 497 GSEITSDVDGDTVLLVAGAPMYFNEGRERKGVYVYELRQNRVYNGTLKDSHYQNA 556

Qy 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFSGSILKTPKQITASELATG 600  
 Db 557 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFSGSILKTPKQITASELATG 616

Qy 601 LQYFGCSIHQGLDNEGLDILAVGALGNVILWSRPVQVNASLHFPSPKINIFHRDCK 660  
 Db 617 LQYFGCSIHQGLDNEGLDILAVGALGNVILWSRPVQVNASLHFPSPKINIFHRDCK 676

Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGI 691  
 Db 677 RSGRDATCLAAFLCFTPIFLAPHFQTTTGI 707

RESULT 10  
 US-10-125-540-313  
 ; Sequence 313, Application US/10125540  
 ; Publication No. US20030059875A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: P7214C1  
 ; CURRENT APPLICATION NUMBER: US/10/125,540  
 ; CURRENT FILING DATE: 2002-04-19  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 646  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 313  
 ; LENGTH: 707  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-125-540-313

Query Match 41.2%; Score 489; DB 14; Length 707;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 11

US-10-108-260A-3386  
 ; Sequence 3386, Application US/10108260A  
 ; Publication No. US20040005560A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HELIX RESEARCH INSTITUTE  
 ; TITLE OF INVENTION: No. US20040005560A1el full length cDNA  
 ; FILE REFERENCE: H1-A0106  
 ; CURRENT APPLICATION NUMBER: US/10/108,260A  
 ; CURRENT FILING DATE: 2002-03-27  
 ; NUMBER OF SEQ ID NOS: 5458  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 3386  
 ; LENGTH: 437  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-108-260A-3386

Query Match 20.6%; Score 245; DB 15; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-225;  
 Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 175 IYPWVEVQHLNLIKFFIYGPQIQGVVQVGEDVWHEFLNDYRSVKDVVEAASHIEQ 234

Db 175 IYFVVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDVVEAASHIEQ 234  
Qy 235 RGTGTTTAFGIEFARSEAFQKGRKGAKKVMIVITGSHSDSPDLEKVIQOQSERDNVTR 294  
Db 235 RGTGTTTAFGIEFARSEAFQKGRKGAKKVMIVITGSHSDSPDLEKVIQOQSERDNVTR 294  
Qy 295 YAVAVLGYNNRGINPFTFLNEIKYIASDPDDKHFFNVTDAAALKDVIDALGDRIFSLEG 354  
Db 295 YAVAVLGYNNRGINPFTFLNEIKYIASDPDDKHFFNVTDAAALKDVIDALGDRIFSLEG 354  
Qy 355 TNKNETSFGLEMSQTGFSHVVEDGVLGAVGAYDMNGAVLKETSAGKVIPLRESYKEF 414  
Db 355 TNKNETSFGLEMSQTGFSHVVEDGVLGAVGAYDMNGAVLKETSAGKVIPLRESYKEF 414  
Qy 415 PEELK 419  
Db 415 PEELK 419

RESULT 12  
US-09-805-354-8  
; Sequence 8, Application US/09805354  
; Publication No. US20030078375A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-536001  
; CURRENT APPLICATION NUMBER: US/09/805,354  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 193  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-805-354-8

Query Match 16.2%; Score 193; DB 10; Length 193;  
Best Local Similarity 100.0%; Pred. No. 1.9e-175;  
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVEHFLND 218  
Db 1 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVEHFLND 60  
Qy 219 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQKGRKGAKKVMIVITGSHSDSP 278  
Db 61 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQKGRKGAKKVMIVITGSHSDSP 120  
Qy 279 DLEKVIQOQSERDNVTRYAVAVLGYNNRGINPFTFLNEIKYIASDPDDKHFFNVTDAAAL 338  
Db 121 DLEKVIQOQSERDNVTRYAVAVLGYNNRGINPFTFLNEIKYIASDPDDKHFFNVTDAAAL 180  
Qy 339 KDVIDALGDRIFS 351  
Db 181 KDVIDALGDRIFS 193

RESULT 13  
US-09-758-493-8  
; Sequence 8, Application US/09758493  
; Publication No. US20040086935A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping

; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: 00786-804001  
; CURRENT APPLICATION NUMBER: US/09/758,493  
; CURRENT FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 193  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-758-493-8

Query Match 16.2%; Score 193; DB 12; Length 193;  
Best Local Similarity 100.0%; Pred. No. 1.9e-175;  
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVEHFLND 218  
Db 1 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVEHFLND 60  
Qy 219 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQKGRKGAKKVMIVITGSHSDSP 278  
Db 61 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQKGRKGAKKVMIVITGSHSDSP 120  
Qy 279 DLEKVIQOQSERDNVTRYAVAVLGYNNRGINPFTFLNEIKYIASDPDDKHFFNVTDAAAL 338  
Db 121 DLEKVIQOQSERDNVTRYAVAVLGYNNRGINPFTFLNEIKYIASDPDDKHFFNVTDAAAL 180  
Qy 339 KDVIDALGDRIFS 351  
Db 181 KDVIDALGDRIFS 193

RESULT 14  
US-10-144-259-8  
; Sequence 8, Application US/10144259  
; Publication No. US20030109691A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-548001  
; CURRENT APPLICATION NUMBER: US/10/144,259  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: US 09/758,493  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 193  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-144-259-8

Query Match 16.2%; Score 193; DB 14; Length 193;  
Best Local Similarity 100.0%; Pred. No. 1.9e-175;  
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVEHFLND 218  
Db 1 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVEHFLND 60  
Qy 219 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQKGRKGAKKVMIVITGSHSDSP 278  
Db 61 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQKGRKGAKKVMIVITGSHSDSP 120  
Qy 279 DLEKVIQOQSERDNVTRYAVAVLGYNNRGINPFTFLNEIKYIASDPDDKHFFNVTDAAAL 338

Db 121 DLEKVIQQSERDNTYAVAVLGYNRRGINPTFLNEIKYIASDPDDKHFENVIDEAL 180  
Qy 339 KOIVDALGDRIFS 351  
Db 181 KOIVDALGDRIFS 193

RESULT 15  
US-09-764-870-472  
; Sequence 472, Application US/09764870  
; Patent No. US20020042386A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT214  
; CURRENT APPLICATION NUMBER: US/09/764,870  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 646  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 472  
; LENGTH: 103  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (96)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-870-472

Query Match 7.4%; Score 88; DB 9; Length 103;  
Best Local Similarity 100.0%; Pred. No. 2.1e-75;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVWHEFHLD 218  
Db 8 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVWHEFHLD 67  
Qy 219 YRSVKDVEAAASHIEQGGTETRTAFGI 246  
Db 68 YRSVKDVEAAASHIEQGGTETRTAFGI 95

RESULT 16  
US-09-764-875-1193  
; Sequence 1193, Application US/09764875  
; Publication No. US20040018969A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P3202  
; CURRENT APPLICATION NUMBER: US/09/764,875  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1249  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1193  
; LENGTH: 103  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (96)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-875-1193

Query Match 7.4%; Score 88; DB 11; Length 103;  
Best Local Similarity 100.0%; Pred. No. 2.1e-75;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVWHEFHLD 218

Db 8 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVWHEFHLD 67  
Qy 219 YRSVKDVEAAASHIEQGGTETRTAFGI 246  
Db 68 YRSVKDVEAAASHIEQGGTETRTAFGI 95

RESULT 17  
US-10-125-540-472  
; Sequence 472, Application US/10125540  
; Publication No. US20030059875A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT214C1  
; CURRENT APPLICATION NUMBER: US/10/125,540  
; CURRENT FILING DATE: 2002-04-19  
; Prior Application removed - See File wrapper or Palm  
; NUMBER OF SEQ ID NOS: 646  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 472  
; LENGTH: 103  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (96)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-125-540-472

Query Match 7.4%; Score 88; DB 14; Length 103;  
Best Local Similarity 100.0%; Pred. No. 2.1e-75;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVWHEFHLD 218  
Db 8 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVWHEFHLD 67  
Qy 219 YRSVKDVEAAASHIEQGGTETRTAFGI 246  
Db 68 YRSVKDVEAAASHIEQGGTETRTAFGI 95

RESULT 18  
US-10-144-259-25  
; Sequence 25, Application US/10144259  
; Publication No. US20030109691A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-548001  
; CURRENT APPLICATION NUMBER: US/10/144,259  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: US 09/758,493  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-144-259-25

Query Match 3.5%; Score 42; DB 14; Length 42;  
Best Local Similarity 100.0%; Pred. No. 7.9e-32;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 335 EAAAKDIVDALGDRIFSLGKTKNKTSPGLEMSQTGFSSHYV 376

Db 1 EAALKDIVDALGDRIFSLGNTKNETSFGLEMSQTGFSSHVV 42

RESULT 19

US-09-866-050A-500

Sequence 500, Application US/09866050A

Publication No. US20030040471A1

GENERAL INFORMATION:

APPLICANT: Watson, James D.

APPLICANT: Strachan, Lorna

APPLICANT: Sleeman, Matthew

APPLICANT: Murison, James G.

APPLICANT: Murison, James G.

APPLICANT: Kumble, Krishanand D.

TITLE OF INVENTION: Compositions Isolated From Skin Cells

TITLE OF INVENTION: and Methods for Their Use

FILE REFERENCE: 11000.1011c4U

CURRENT APPLICATION NUMBER: US/09/866.050A

CURRENT FILING DATE: 2001-05-24

NUMBER OF SEQ ID NOS: 725

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 500

LENGTH: 545

TYPE: PRT

ORGANISM: Mouse

US-09-866-050A-500

Query Match 3.5%; Score 42; DB 10; Length 545;

Best Local Similarity 100.0%; Pred. No. 7.8e-31;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 GWPTTLRVSVFPWNGCNEDEHCVPDLVLDARSDLPTAMEYCO 807

Db 123 GWPTTLRVSVFPWNGCNEDEHCVPDLVLDARSDLPTAMEYCO 164

RESULT 20

US-09-866-050A-624

Sequence 624, Application US/09866050A

Publication No. US20030040471A1

GENERAL INFORMATION:

APPLICANT: Watson, James D.

APPLICANT: Strachan, Lorna

APPLICANT: Sleeman, Matthew

APPLICANT: Murison, James G.

APPLICANT: Murison, James G.

APPLICANT: Kumble, Krishanand D.

TITLE OF INVENTION: Compositions Isolated From Skin Cells

TITLE OF INVENTION: and Methods for Their Use

FILE REFERENCE: 11000.1011c4U

CURRENT APPLICATION NUMBER: US/09/866.050A

CURRENT FILING DATE: 2001-05-24

NUMBER OF SEQ ID NOS: 725

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 624

LENGTH: 688

TYPE: PRT

ORGANISM: Rat

US-09-866-050A-624

Query Match 3.5%; Score 42; DB 10; Length 688;

Best Local Similarity 100.0%; Pred. No. 9.6e-31;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 GWPTTLRVSVFPWNGCNEDEHCVPDLVLDARSDLPTAMEYCO 807

Db 266 GWPTTLRVSVFPWNGCNEDEHCVPDLVLDARSDLPTAMEYCO 307

RESULT 21

US-09-866-050A-501

Sequence 501, Application US/09866050A

Publication No. US20030040471A1

GENERAL INFORMATION:

APPLICANT: Watson, James D.

APPLICANT: Strachan, Lorna

APPLICANT: Sleeman, Matthew

APPLICANT: Murison, James G.

APPLICANT: Murison, James G.

APPLICANT: Kumble, Krishanand D.

TITLE OF INVENTION: Compositions Isolated From Skin Cells

TITLE OF INVENTION: and Methods for Their Use

FILE REFERENCE: 11000.1011c4U

CURRENT APPLICATION NUMBER: US/09/866.050A

CURRENT FILING DATE: 2001-05-24

NUMBER OF SEQ ID NOS: 725

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 501

LENGTH: 696

TYPE: PRT

ORGANISM: Rat

US-09-866-050A-501

Query Match 3.5%; Score 42; DB 10; Length 696;

Best Local Similarity 100.0%; Pred. No. 9.7e-31;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 GWPTTLRVSVFPWNGCNEDEHCVPDLVLDARSDLPTAMEYCO 807

Db 266 GWPTTLRVSVFPWNGCNEDEHCVPDLVLDARSDLPTAMEYCO 307

RESULT 22

US-09-984-130-102

Sequence 102, Application US/09984130

Publication No. US2003005231A1

GENERAL INFORMATION:

APPLICANT: Ni et al.

TITLE OF INVENTION: 12 Human Secreted Proteins

FILE REFERENCE: PF489P2

CURRENT APPLICATION NUMBER: US/09/984.130

CURRENT FILING DATE: 2001-10-29

PRIOR APPLICATION NUMBER: 60/243,792

PRIOR FILING DATE: 2000-10-30

PRIOR APPLICATION NUMBER: 09/836,353

PRIOR FILING DATE: 2001-04-18

PRIOR APPLICATION NUMBER: 60/198,407

PRIOR FILING DATE: 2000-04-19

PRIOR APPLICATION NUMBER: PCT/US99/25031

PRIOR FILING DATE: 1999-10-27

PRIOR APPLICATION NUMBER: 60/105,971

PRIOR FILING DATE: 1998-10-28

NUMBER OF SEQ ID NOS: 149

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 102

LENGTH: 33

TYPE: PRT

ORGANISM: Homo sapiens

US-09-984-130-102

Query Match 2.8%; Score 33; DB 10; Length 33;

Best Local Similarity 100.0%; Pred. No. 2.5e-23;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 64 TNGYQKTGVDYKCPVTHGNCCTKLNLRVTLSNV 96

Db 1 TNGYQKTGVDYKCPVTHGNCCTKLNLRVTLSNV 33

RESULT 23

US-09-836-353A-102

Sequence 102, Application US/09836353A

Publication No. US20030129685A1

GENERAL INFORMATION:

APPLICANT: Ni et al.

TITLE OF INVENTION: 12 Human Secreted Proteins

FILE REFERENCE: PF489P2

CURRENT APPLICATION NUMBER: US/09/836.353

CURRENT FILING DATE: 2001-10-29

PRIOR APPLICATION NUMBER: 60/243,792

PRIOR FILING DATE: 2000-10-30

PRIOR APPLICATION NUMBER: 09/836,353

PRIOR FILING DATE: 2001-04-18

PRIOR APPLICATION NUMBER: 60/198,407

PRIOR FILING DATE: 2000-04-19

PRIOR APPLICATION NUMBER: PCT/US99/25031

PRIOR FILING DATE: 1999-10-27

PRIOR APPLICATION NUMBER: 60/105,971

PRIOR FILING DATE: 1998-10-28

NUMBER OF SEQ ID NOS: 149

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 102

LENGTH: 33

TYPE: PRT

ORGANISM: Homo sapiens

US-09-836-353A-102

Query Match 2.8%; Score 33; DB 10; Length 33;

Best Local Similarity 100.0%; Pred. No. 2.5e-23;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 64 TNGYQKTGVDYKCPVTHGNCCTKLNLRVTLSNV 96

Db 1 TNGYQKTGVDYKCPVTHGNCCTKLNLRVTLSNV 33

```
; FILE REFERENCE: PF489P1
; CURRENT APPLICATION NUMBER: US/09/836,353A
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; SOFTWARE: PatentIn Ver. 2.0
; NUMBER OF SEQ ID NOS: 147
; SEQ ID NO 102
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-353A-102

Query Match      2.8%; Score 33; DB 10; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.5e-23;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 TNGYQKTGVDYKCPVHGNCTKLNLRVTLNV 96
Db 1 TNGYQKTGVDYKCPVHGNCTKLNLRVTLNV 33

RESULT 24
US-09-866-050A-391
; Sequence 391, Application US/09866050A
; Publication No. US2003004071A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 391
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-391

Query Match      1.8%; Score 21; DB 10; Length 158;
Best Local Similarity 100.0%; Pred. No. 2.9e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 RPVVQINASLHPEPSKINIFH 656
Db 57 RPVVQINASLHPEPSKINIFH 77

RESULT 25
US-09-866-050A-390
; Sequence 390, Application US/09866050A
; Publication No. US2003004071A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
```

```
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 390
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-390

Query Match      1.4%; Score 17; DB 10; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 NGTLKDSHSYQNAFPGS 544
Db 28 NGTLKDSHSYQNAFPGS 44

RESULT 26
US-09-764-870-284
; Sequence 284, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 284
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-284

Query Match      1.3%; Score 15; DB 9; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYFW 178
Db 47 DIVIVLDGNSIYFW 61

RESULT 27
US-09-764-870-444
; Sequence 444, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 444
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-444

Query Match      1.3%; Score 15; DB 9; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYFW 178
Db 47 DIVIVLDGNSIYFW 61
```

## RESULT 28

US-10-125-540-284  
; Sequence 284, Application US/10125540  
; Publication No. US20030059875A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT214C1  
; CURRENT APPLICATION NUMBER: US/10/125,540  
; CURRENT FILING DATE: 2002-04-19  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 646  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 284  
; LENGTH: 148  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-125-540-284

Query Match 1.3%; Score 15; DB 14; Length 148;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGSGNSIYPW 178  
Db 47 DIVVLGSGNSIYPW 61

## RESULT 29

US-10-125-540-444  
; Sequence 444, Application US/10125540  
; Publication No. US20030059875A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT214C1  
; CURRENT APPLICATION NUMBER: US/10/125,540  
; CURRENT FILING DATE: 2002-04-19  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 646  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 444  
; LENGTH: 148  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-125-540-444

Query Match 1.3%; Score 15; DB 14; Length 148;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGSGNSIYPW 178  
Db 47 DIVVLGSGNSIYPW 61

## RESULT 30

US-10-346-863-57  
; Sequence 57, Application US/10346863  
; Publication No. US20040038325A1  
; GENERAL INFORMATION:  
; APPLICANT: PHELPS, CHRISTOPHER BENJAMIN  
; APPLICANT: FAGAN, RICHARD JOSEPH  
; APPLICANT: GUTTERIDGE, ALEX  
; TITLE OF INVENTION: ADHESION MOLECULES  
; FILE REFERENCE: 674575-2001  
; CURRENT APPLICATION NUMBER: US/10/346,863  
; CURRENT FILING DATE: 2003-01-17  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 57  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-346-863-57

Query Match 1.3%; Score 15; DB 12; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGSGNSIYPW 178  
Db 5 DIVVLGSGNSIYPW 19

## RESULT 31

US-10-474-832-59  
; Sequence 59, Application US/10474832  
; Publication No. US20040081651A1  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; TITLE OF INVENTION: ANTIBODIES TO VLA-1  
; FILE REFERENCE: A101 PCT  
; CURRENT APPLICATION NUMBER: US/10/474,832  
; CURRENT FILING DATE: 2003-10-14  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 59  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Rat/human chimeric  
; OTHER INFORMATION: I domain construct  
US-10-474-832-59

Query Match 1.3%; Score 15; DB 16; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGSGNSIYPW 178  
Db 4 DIVVLGSGNSIYPW 18

## RESULT 32

US-10-474-832-60  
; Sequence 60, Application US/10474832  
; Publication No. US20040081651A1  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; TITLE OF INVENTION: ANTIBODIES TO VLA-1  
; FILE REFERENCE: A101 PCT  
; CURRENT APPLICATION NUMBER: US/10/474,832  
; CURRENT FILING DATE: 2003-10-14  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 60  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-474-832-60

Query Match 1.3%; Score 15; DB 12; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGSGNSIYPW 178  
Db 5 DIVVLGSGNSIYPW 19



; ORGANISM: Rattus sp.  
US-10-474-832-60

Query Match 1.3%; Score 15; DB 16; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178  
|||||  
DB 4 DIVVLGSGNSIYPW 18

RESULT 33

US-10-474-832-61  
; Sequence 61, Application US/10474832  
; Publication No. US20040081651A1  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; TITLE OF INVENTION: ANTIBODIES TO VLA-1  
; FILE REFERENCE: A101 PCT  
; CURRENT APPLICATION NUMBER: US/10/474,832  
; CURRENT FILING DATE: 2003-10-14  
; PRIOR APPLICATION NUMBER: 60/283,794  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: 60/303,689  
; PRIOR FILING DATE: 2001-07-06  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 61  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-474-832-61

Query Match 1.3%; Score 15; DB 16; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178  
|||||  
DB 4 DIVVLGSGNSIYPW 18

RESULT 34

US-09-805-354-5  
; Sequence 5, Application US/09805354  
; Publication No. US20030078375A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-536001  
; CURRENT APPLICATION NUMBER: US/09/805,354  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: US 09/758,493  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-805-354-5

Query Match 1.3%; Score 15; DB 10; Length 195;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178  
|||||

Db 6 DIVVLGSGNSIYPW 20

RESULT 35

US-09-758-493-5  
; Sequence 5, Application US/09758493  
; Publication No. US20040086935A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-804001  
; CURRENT APPLICATION NUMBER: US/09/758,493  
; CURRENT FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-758-493-5

Query Match 1.3%; Score 15; DB 12; Length 195;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178  
|||||  
DB 6 DIVVLGSGNSIYPW 20

RESULT 36

US-10-144-259-5  
; Sequence 5, Application US/10144259  
; Publication No. US20030109691A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-548001  
; CURRENT APPLICATION NUMBER: US/10/144,259  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: US 09/758,493  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-144-259-5

Query Match 1.3%; Score 15; DB 14; Length 195;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178  
|||||  
DB 6 DIVVLGSGNSIYPW 20

RESULT 37

US-09-996-738-5  
; Sequence 5, Application US/09996738  
; Patent No. US20020146417A1  
; GENERAL INFORMATION:

```

; APPLICANT: Biogen, Inc.
; APPLICANT: De Fougères, Antonin
; APPLICANT: Gotwals, Philip
; APPLICANT: Lobb, Roy
; APPLICANT: Koteliansky, Victor
; TITLE OF INVENTION: Method for the Treatment of Inflammatory
; TITLE OF INVENTION: Disorders
; FILE REFERENCE: A076PCT
; CURRENT APPLICATION NUMBER: US/09/996,738
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/185336
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 60/137038
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 212
; TYPE: PRT
; ORGANISM: rat
US-09-996-738-5

```

```

Query Match      1.3%; Score 15; DB 9; Length 212;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 164 DIVVLGSGNSIYPW 178
Db 22 DIVVLGSGNSIYPW 36

```

```

RESULT 38
US-09-996-738-5
; Sequence 6, Application US/09996738
; Patent No. US20020146417A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: De Fougères, Antonin
; APPLICANT: Gotwals, Philip
; APPLICANT: Lobb, Roy
; APPLICANT: Koteliansky, Victor
; TITLE OF INVENTION: Method for the Treatment of Inflammatory
; TITLE OF INVENTION: Disorders
; FILE REFERENCE: A076PCT
; CURRENT APPLICATION NUMBER: US/09/996,738
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/185336
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 60/137038
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-996-738-6

```

```

Query Match      1.3%; Score 15; DB 9; Length 212;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 164 DIVVLGSGNSIYPW 178
Db 22 DIVVLGSGNSIYPW 36

```

```

RESULT 39
US-10-625-260-5
; Sequence 5, Application US/10625260
; Publication No. US20040037827A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.

```

```

; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliansky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/625,260
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Rat
US-10-625-260-5

```

```

Query Match      1.3%; Score 15; DB 12; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 164 DIVVLGSGNSIYPW 178
Db 22 DIVVLGSGNSIYPW 36

```

```

RESULT 40
US-10-625-260-6
; Sequence 6, Application US/10625260
; Publication No. US20040037827A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliansky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/625,260
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-625-260-6

```

```

Query Match      1.3%; Score 15; DB 12; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 164 DIVVLGSGNSIYPW 178
Db 22 DIVVLGSGNSIYPW 36

```

```

RESULT 41
US-10-625-260-9
; Sequence 9, Application US/10625260
; Publication No. US20040037827A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliansky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/625,260
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: 60/137,214

```

```

; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-625-260-9

```

```

Query Match      1.3%: Score 15; DB 12; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      164 DIVVLDGNSIYPW 178
Db      22 DIVVLDGNSIYPW 36

```

## RESULT 42

```

US-10-061-658-5
; Sequence 5, Application US/10061658
; Publication No. US20020182213A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A07305
; CURRENT APPLICATION NUMBER: US/10/061,658
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Rat
US-10-061-658-5

```

```

Query Match      1.3%: Score 15; DB 13; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      164 DIVVLDGNSIYPW 178
Db      22 DIVVLDGNSIYPW 36

```

## RESULT 43

```

US-10-061-658-6
; Sequence 6, Application US/10061658
; Publication No. US20020182213A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A07305
; CURRENT APPLICATION NUMBER: US/10/061,658
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 214

```

```

; TYPE: PRT
; ORGANISM: Homo sapien
US-10-061-658-6

```

```

Query Match      1.3%: Score 15; DB 13; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      164 DIVVLDGNSIYPW 178
Db      22 DIVVLDGNSIYPW 36

```

## RESULT 44

```

US-10-061-658-9
; Sequence 9, Application US/10061658
; Publication No. US20020182213A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A07305
; CURRENT APPLICATION NUMBER: US/10/061,658
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-061-658-9

```

```

Query Match      1.3%: Score 15; DB 13; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      164 DIVVLDGNSIYPW 178
Db      22 DIVVLDGNSIYPW 36

```

## RESULT 45

```

US-10-474-832-63
; Sequence 63, Application US/10474832
; Publication No. US20040081651A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: ANTIBODIES TO VLA-1
; FILE REFERENCE: A101 PCT
; CURRENT APPLICATION NUMBER: US/10/474,832
; CURRENT FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/283,794
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/303,689
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-474-832-63

```

```

Query Match      1.3%: Score 15; DB 16; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      164 DIVVLDGNSIYPW 178

```

```
Db      22 DIVVLGSGNSIYPW 36
|||||
RESULT 46
US-10-474-832-64
; Sequence 64, Application US/10474832
; Publication No. US20040081651A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: ANTIBODIES TO VLA-1
; FILE REFERENCE: A101 PCT
; CURRENT APPLICATION NUMBER: US/10/474,832
; CURRENT FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/283,794
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/303,689
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-474-832-64

Query Match      1.3%; Score 15; DB 16; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVVLGSGNSIYPW 178
|||||
Db      22 DIVVLGSGNSIYPW 36
|||||

RESULT 47
US-09-984-130-103
; Sequence 103, Application US/09984130
; Publication No. US20030055231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: P489P2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 103
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-103

Query Match      1.3%; Score 15; DB 10; Length 1151;
Best Local Similarity 100.0%; Pred. No. 9.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVVLGSGNSIYPW 178
|||||
Db      144 DIVVLGSGNSIYPW 158
|||||

RESULT 48
US-09-984-130-103
; Sequence 103, Application US/09836353A
; Publication No. US20030129685A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: P489P1
; CURRENT APPLICATION NUMBER: US/09/836,353A
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 103
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-353A-103

Query Match      1.3%; Score 15; DB 10; Length 1151;
Best Local Similarity 100.0%; Pred. No. 9.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVVLGSGNSIYPW 178
|||||
Db      144 DIVVLGSGNSIYPW 158
|||||

RESULT 49
US-09-918-715-250
; Sequence 250, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Brad St. Croix
; APPLICANT: Bert Vogelstein
; APPLICANT: Kenneth Kinzler
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00134
; CURRENT APPLICATION NUMBER: US/09/918,715
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/222,599
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/224,360
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 250
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-918-715-250

Query Match      1.3%; Score 15; DB 12; Length 1179;
Best Local Similarity 100.0%; Pred. No. 9.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVVLGSGNSIYPW 178
|||||
Db      172 DIVVLGSGNSIYPW 186
|||||

RESULT 50
US-09-918-715-307
; Sequence 307, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Brad St. Croix
```

; APPLICANT: Bert Vogelstein  
; APPLICANT: Kenneth Kinzler  
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS  
; FILE REFERENCE: 1107.00134  
; CURRENT APPLICATION NUMBER: US/09/918,715  
; CURRENT FILING DATE: 2001-08-01  
; PRIOR APPLICATION NUMBER: 60/222,599  
; PRIOR FILING DATE: 2000-08-02  
; PRIOR APPLICATION NUMBER: 60/224,360  
; PRIOR FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: 60/282,850  
; PRIOR FILING DATE: 2000-04-11  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 307  
; LENGTH: 1180  
; TYPE: PRT  
; ORGANISM: Rat  
US-09-918-715-307

Query Match 1.1%; Score 15; DB 12; Length 1180;  
Best Local Similarity 100.0%; Pred. No. 9.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178  
DB 172 DIVVLGSGNSIYPW 186  
|||||

RESULT 51  
US-09-805-354-7  
; Sequence 7, Application US/09805354  
; Publication No. US20030078375A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-536001  
; CURRENT APPLICATION NUMBER: US/09/805,354  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: US 09/758,493  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-805-354-7

Query Match 1.1%; Score 13; DB 10; Length 195;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 VIVLDGSGNSIYPW 178  
DB 8 VIVLDGSGNSIYPW 20  
|||||

RESULT 52  
US-09-758-493-7  
; Sequence 7, Application US/09758493  
; Publication No. US20040086935A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND  
; USES THEREOF  
; FILE REFERENCE: 00786-804001

; CURRENT APPLICATION NUMBER: US/09/758,493  
; CURRENT FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-758-493-7

Query Match 1.1%; Score 13; DB 12; Length 195;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 VIVLDGSGNSIYPW 178  
DB 8 VIVLDGSGNSIYPW 20  
|||||

RESULT 53  
US-10-144-259-7  
; Sequence 7, Application US/10144259  
; Publication No. US20030109691A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-548001  
; CURRENT APPLICATION NUMBER: US/10/144,259  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: US 09/758,493  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-144-259-7

Query Match 1.1%; Score 13; DB 14; Length 195;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 VIVLDGSGNSIYPW 178  
DB 8 VIVLDGSGNSIYPW 20  
|||||

RESULT 54  
US-10-144-259-26  
; Sequence 26, Application US/10144259  
; Publication No. US20030109691A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-548001  
; CURRENT APPLICATION NUMBER: US/10/144,259  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: US 09/758,493  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 26

```
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-26

Query Match          0.9%; Score 11; DB 14; Length 43;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  340 DIVDALGORIF 350
    |||||
Db   6 DIVDALGDRIF 16

RESULT 55
US-10-336-603A-42
; Sequence 42, Application US/10336603A
; Publication No. US20040072997A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-533A
; CURRENT APPLICATION NUMBER: US/10/336,603A
; CURRENT FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: 09/746,491
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 10/055,569
; PRIOR FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 42
; LENGTH: 1147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-336-603A-42

Query Match          0.9%; Score 11; DB 12; Length 1147;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  495 TDVLLVGAPMY 505
    |||||
Db   506 TDVLLVGAPMY 516

RESULT 56
US-10-211-462-187
; Sequence 187, Application US/10211462
; Publication No. US20040033495A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 187
; LENGTH: 1181
; TYPE: PRT

; ORGANISM: Homo sapiens
US-10-211-462-187

Query Match          0.9%; Score 11; DB 12; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  495 TDVLLVGAPMY 505
    |||||
Db   506 TDVLLVGAPMY 516

RESULT 57
US-10-160-354-2
; Sequence 2, Application US/10160354
; Publication No. US20030013107A1
; GENERAL INFORMATION:
; APPLICANT: Chu, Peter
; APPLICANT: Li, Congfen
; APPLICANT: Liao, X. Charlene
; APPLICANT: Pardo, Jorge
; APPLICANT: Rigol Pharmaceuticals, Incorporated
; TITLE OF INVENTION: Alpha 2 Intergrin: Modulators of Lymphocyte Activation
; FILE REFERENCE: 021044-001110US
; CURRENT APPLICATION NUMBER: US/10/160,354
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: US 60/296,819
; PRIOR FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1181
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-160-354-2

Query Match          0.9%; Score 11; DB 14; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  495 TDVLLVGAPMY 505
    |||||
Db   506 TDVLLVGAPMY 516

RESULT 58
US-10-295-027-1286
; Sequence 1286, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
```

; PRIOR APPLICATION NUMBER: US 60/332,454  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 60/334,393  
; PRIOR FILING DATE: 2001-11-29  
; PRIOR APPLICATION NUMBER: US 60/340,376  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: US 60/347,211  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 60/347,349  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/355,250  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: US 60/356,714  
; PRIOR FILING DATE: 2002-02-13  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1386  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1286  
; LENGTH: 1181  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-295-027-1286

Query Match 0.8%; Score 11; DB 15; Length 1181;  
Best Local Similarity 100.0%; Pred. No. 0.63; Mismatches 0; Indels 0; Gaps 0;  
Matches 11; Conservative 0;

Qy 495 TDVLLVGAPMY 505  
|||  
Db 506 TDVLLVGAPMY 516

## RESULT 59

US-10-108-260A-3415  
; Sequence 3415, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA  
; FILE REFERENCE: HI-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3415  
; LENGTH: 164  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-3415

Query Match 0.8%; Score 9; DB 15; Length 164;  
Best Local Similarity 100.0%; Pred. No. 8.8; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

Qy 1153 LLLALLVL 1161  
|||  
Db 99 LLLALLVL 107

## RESULT 60

US-10-038-854-102  
; Sequence 102, Application US/10038854  
; Publication No. US20040022781A1  
; GENERAL INFORMATION:  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Li, Li  
; APPLICANT: Wolenc, Adam R  
; APPLICANT: Vernet, Corine  
; APPLICANT: Eisen, Andrew J  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Tchernev, Velizar

; APPLICANT: Spaderna, Steven K  
; APPLICANT: Gorman, Linda  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Gusev, Vladimir Y  
; APPLICANT: Gangolli, Esha A  
; APPLICANT: Guo, Xiaojia S  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Bológ, Ferenc  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Gunther, Erik  
; APPLICANT: Smithson, Glennada  
; APPLICANT: Millet, Isabelle  
; APPLICANT: MacDougall, John R  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-230  
; CURRENT APPLICATION NUMBER: US/10/038,854  
; CURRENT FILING DATE: 2003-01-22  
; PRIOR APPLICATION NUMBER: 60/258,928  
; PRIOR FILING DATE: 2000-12-29  
; PRIOR APPLICATION NUMBER: 60/259,415  
; PRIOR FILING DATE: 2001-01-02  
; PRIOR APPLICATION NUMBER: 60/259,785  
; PRIOR FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: 60/269,814  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 60/279,832  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/279,833  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/279,863  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/283,889  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: 60/284,447  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/286,683  
; PRIOR FILING DATE: 2001-04-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 411  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 102  
; LENGTH: 312  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-10-038-854-102

Query Match 0.8%; Score 9; DB 16; Length 312;  
Best Local Similarity 100.0%; Pred. No. 16; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

Qy 1154 LLLALLVL 1162  
|||  
Db 3 LLLALLVL 11

## RESULT 61

US-10-369-493-2470  
; Sequence 2470, Application US/10369493  
; Publication No. US2003023675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B



; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 2470  
 ; LENGTH: 371  
 ; TYPE: PRT  
 ; ORGANISM: Schizosaccharomyces pombe  
 US-10-369-493-2470

Query Match 0.8%; Score 9; DB 15; Length 371;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 DGVLLGAVG 386  
 Db 69 DGVLLGAVG 77  
 |||||

## RESULT 62

US-09-350-259-20  
 ; Sequence 20, Application US/09350259  
 ; Patent No. US20020062008A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gallatin, Michael W.  
 ; APPLICANT: Van der Vieren, Monica  
 ; TITLE OF INVENTION: No. US20020062008A1el Human 2  
 ; FILE REFERENCE: 27866/35004  
 ; CURRENT APPLICATION NUMBER: US/09/350,259  
 ; CURRENT FILING DATE: 1999-07-08  
 ; EARLIER APPLICATION NUMBER: 09/193,043  
 ; EARLIER FILING DATE: 1998-11-16  
 ; EARLIER APPLICATION NUMBER: 08/173,497  
 ; EARLIER FILING DATE: 1993-12-23  
 ; EARLIER APPLICATION NUMBER: 08/286,889  
 ; EARLIER FILING DATE: 1994-08-05  
 ; EARLIER APPLICATION NUMBER: 08/362,852  
 ; EARLIER FILING DATE: 1994-12-21  
 ; EARLIER APPLICATION NUMBER: 08/943,363  
 ; EARLIER FILING DATE: 1997-10-03  
 ; NUMBER OF SEQ ID NOS: 114  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 20  
 ; LENGTH: 17  
 ; TYPE: PRT  
 ; ORGANISM: dog  
 US-09-350-259-20

Query Match 0.7%; Score 8; DB 9; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
 Db 1 LVVGAPLE 8  
 |||||

## RESULT 63

US-09-891-943-20  
 ; Sequence 20, Application US/09891943  
 ; Publication No. US2003007278A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gallatin, Michael W.  
 ; APPLICANT: Van der Vieren, Monica  
 ; TITLE OF INVENTION: No. US2003007278A1el Human 2  
 ; FILE REFERENCE: 27866/35004  
 ; CURRENT APPLICATION NUMBER: US/09/891,943  
 ; CURRENT FILING DATE: 2001-06-26  
 ; PRIOR APPLICATION NUMBER: 09/193,043  
 ; PRIOR FILING DATE: 1998-11-16  
 ; PRIOR APPLICATION NUMBER: 08/286,889  
 ; PRIOR FILING DATE: 1994-08-05

; PRIOR APPLICATION NUMBER: 08/362,652  
 ; PRIOR FILING DATE: 1994-12-21  
 ; PRIOR APPLICATION NUMBER: 08/943,363  
 ; PRIOR FILING DATE: 1997-10-03  
 ; NUMBER OF SEQ ID NOS: 114  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 20  
 ; LENGTH: 17  
 ; TYPE: PRT  
 ; ORGANISM: dog  
 US-09-891-943-20

Query Match 0.7%; Score 8; DB 10; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
 Db 1 LVVGAPLE 8  
 |||||

## RESULT 64

US-10-144-259-27  
 ; Sequence 27, Application US/10144259  
 ; Publication No. US20030109691A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Arnaout, M. Amin  
 ; APPLICANT: Li, Rui  
 ; APPLICANT: Xiong, Jian-Ping  
 ; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF  
 ; FILE REFERENCE: 00786-548001  
 ; CURRENT APPLICATION NUMBER: US/10/144,259  
 ; CURRENT FILING DATE: 2002-09-04  
 ; PRIOR APPLICATION NUMBER: US 09/758,493  
 ; PRIOR FILING DATE: 2001-01-11  
 ; PRIOR APPLICATION NUMBER: US 60/221,950  
 ; PRIOR FILING DATE: 2000-07-31  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 27  
 ; LENGTH: 43  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-144-259-27

Query Match 0.7%; Score 8; DB 14; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 365 EMSQTGFS 372  
 Db 32 EMSQTGFS 39  
 |||||

## RESULT 65

US-09-978-295A-85  
 ; Sequence 85, Application US/09978295A  
 ; Patent No. US20020156006A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Baker Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnovers, Luc  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Ferrara, Napoleon  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gertsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gurney, Austin L.  
 APPLICANT: Hillan, Kenneth J.  
 APPLICANT: Kijavio, Ivar J.  
 APPLICANT: Kuo, Sophia S.  
 APPLICANT: Napier, Mary A.  
 APPLICANT: Pan, James.  
 APPLICANT: Paoni, Nicholas F.  
 APPLICANT: Roy, Margaret Ann  
 APPLICANT: Shelton, David L.  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Williams, P. Micky  
 APPLICANT: Wood, William I.  
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 FILE OF INVENTION: Acids Encoding the Same  
 FILE REFERENCE: P2630P1C11  
 CURRENT APPLICATION NUMBER: US/09/978,295A  
 CURRENT FILING DATE: 2001-10-15  
 PRIOR APPLICATION NUMBER: 09/518585  
 PRIOR FILING DATE: 2001-07-30  
 PRIOR APPLICATION NUMBER: 60/062250  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/064249  
 PRIOR FILING DATE: 1997-11-03  
 PRIOR APPLICATION NUMBER: 60/065311  
 PRIOR FILING DATE: 1997-11-13  
 PRIOR APPLICATION NUMBER: 60/066364  
 PRIOR FILING DATE: 1997-11-21  
 PRIOR APPLICATION NUMBER: 60/077450  
 PRIOR FILING DATE: 1998-03-10  
 PRIOR APPLICATION NUMBER: 60/077632  
 PRIOR FILING DATE: 1998-03-11  
 PRIOR APPLICATION NUMBER: 60/077641  
 PRIOR FILING DATE: 1998-03-11  
 PRIOR APPLICATION NUMBER: 60/077649  
 PRIOR FILING DATE: 1998-03-11  
 PRIOR APPLICATION NUMBER: 60/077791  
 PRIOR FILING DATE: 1998-03-12  
 PRIOR APPLICATION NUMBER: 60/078004  
 PRIOR FILING DATE: 1998-03-13  
 PRIOR APPLICATION NUMBER: 60/078886  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/078936  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/078939  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/079294  
 PRIOR FILING DATE: 1998-03-25  
 PRIOR APPLICATION NUMBER: 60/079656  
 PRIOR FILING DATE: 1998-03-26  
 PRIOR APPLICATION NUMBER: 60/079664  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079689  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079663  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079728  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079786  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079920  
 PRIOR FILING DATE: 1998-03-30  
 PRIOR APPLICATION NUMBER: 60/079923  
 PRIOR FILING DATE: 1998-03-30  
 PRIOR APPLICATION NUMBER: 60/080105  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080107  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080165  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080194

PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080327  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080328  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080333  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080334  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/081070  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081049  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081071  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081195  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081203  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081229  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081955  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081817  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081819  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081952  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081838  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/082568  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082569  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082704  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082804  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082700  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082797  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082796  
 PRIOR FILING DATE: 1998-04-23  
 PRIOR APPLICATION NUMBER: 60/083336  
 PRIOR FILING DATE: 1998-04-27  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/083392  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083495  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083496  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083499  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083545  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083554  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083558  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083559  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083500  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083742  
 PRIOR FILING DATE: 1998-04-30  
 PRIOR APPLICATION NUMBER: 60/084366  
 PRIOR FILING DATE: 1998-05-05  
 PRIOR APPLICATION NUMBER: 60/084414  
 PRIOR FILING DATE: 1998-05-06

```

; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

```

Query Match 0.7%; Score 8; DB 9; Length 67;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1152 GLLIALL 1159
Db 12 GLLIALL 19

```

```

RESULT 66
US-09-978-697-85
; Sequence 85, Application US/09978697
; Patent No. US2002016928A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Eaton, Dan
; APPLICANT: Ferraro, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gottard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas P.

```

```

; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC27
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01

```

PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640

PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084643  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085689  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085573  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 9; Length 67;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159  
DB 12 GLLLLALL 19

RESULT 67

US-09-978-192A-85  
Sequence 85, Application US/09978192A  
Patent No. US2002017753A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Quang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Garney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

1 TITLE OF INVENTION: Acids Encoding the Same  
2 FILE REFERENCE: 22630PIC9  
3 CURRENT APPLICATION NUMBER: US/09/978,192A  
4 CURRENT FILING DATE: 2001-10-15  
5 PRIOR APPLICATION NUMBER: 09/918585  
6 PRIOR FILING DATE: 2001-07-30  
7 PRIOR APPLICATION NUMBER: 60/062250  
8 PRIOR FILING DATE: 1997-10-17  
9 PRIOR APPLICATION NUMBER: 60/064249  
10 PRIOR FILING DATE: 1997-11-03  
11 PRIOR APPLICATION NUMBER: 60/065311  
12 PRIOR FILING DATE: 1997-11-13  
13 PRIOR APPLICATION NUMBER: 60/066364  
14 PRIOR FILING DATE: 1997-11-21  
15 PRIOR APPLICATION NUMBER: 60/077450  
16 PRIOR FILING DATE: 1998-03-10  
17 PRIOR APPLICATION NUMBER: 60/077632  
18 PRIOR FILING DATE: 1998-03-11  
19 PRIOR APPLICATION NUMBER: 60/077641  
20 PRIOR FILING DATE: 1998-03-11  
21 PRIOR APPLICATION NUMBER: 60/077649  
22 PRIOR FILING DATE: 1998-03-11  
23 PRIOR APPLICATION NUMBER: 60/077791  
24 PRIOR FILING DATE: 1998-03-12  
25 PRIOR APPLICATION NUMBER: 60/078004  
26 PRIOR FILING DATE: 1998-03-13  
27 PRIOR APPLICATION NUMBER: 60/078886  
28 PRIOR FILING DATE: 1998-03-20  
29 PRIOR APPLICATION NUMBER: 60/078936  
30 PRIOR FILING DATE: 1998-03-20  
31 PRIOR APPLICATION NUMBER: 60/078910  
32 PRIOR FILING DATE: 1998-03-20  
33 PRIOR APPLICATION NUMBER: 60/078939  
34 PRIOR FILING DATE: 1998-03-20  
35 PRIOR APPLICATION NUMBER: 60/079294  
36 PRIOR FILING DATE: 1998-03-25  
37 PRIOR APPLICATION NUMBER: 60/079656  
38 PRIOR FILING DATE: 1998-03-26  
39 PRIOR APPLICATION NUMBER: 60/079664  
40 PRIOR FILING DATE: 1998-03-27  
41 PRIOR APPLICATION NUMBER: 60/079689  
42 PRIOR FILING DATE: 1998-03-27  
43 PRIOR APPLICATION NUMBER: 60/079663  
44 PRIOR FILING DATE: 1998-03-27  
45 PRIOR APPLICATION NUMBER: 60/079728  
46 PRIOR FILING DATE: 1998-03-27  
47 PRIOR APPLICATION NUMBER: 60/079786  
48 PRIOR FILING DATE: 1998-03-27  
49 PRIOR APPLICATION NUMBER: 60/079920  
50 PRIOR FILING DATE: 1998-03-30  
51 PRIOR APPLICATION NUMBER: 60/079923  
52 PRIOR FILING DATE: 1998-03-30  
53 PRIOR APPLICATION NUMBER: 60/080105  
54 PRIOR FILING DATE: 1998-03-31  
55 PRIOR APPLICATION NUMBER: 60/080107  
56 PRIOR FILING DATE: 1998-03-31  
57 PRIOR APPLICATION NUMBER: 60/080165  
58 PRIOR FILING DATE: 1998-03-31  
59 PRIOR APPLICATION NUMBER: 60/080194  
60 PRIOR FILING DATE: 1998-03-31  
61 PRIOR APPLICATION NUMBER: 60/080327  
62 PRIOR FILING DATE: 1998-04-01  
63 PRIOR APPLICATION NUMBER: 60/080328  
64 PRIOR FILING DATE: 1998-04-01  
65 PRIOR APPLICATION NUMBER: 60/080333  
66 PRIOR FILING DATE: 1998-04-01  
67 PRIOR APPLICATION NUMBER: 60/080334  
68 PRIOR FILING DATE: 1998-04-01  
69 PRIOR APPLICATION NUMBER: 60/081070  
70 PRIOR FILING DATE: 1998-04-08  
71 PRIOR APPLICATION NUMBER: 60/081049  
72 PRIOR FILING DATE: 1998-04-08  
73 PRIOR APPLICATION NUMBER: 60/081071  
74 PRIOR FILING DATE: 1998-04-08  
75 PRIOR APPLICATION NUMBER: 60/081195  
76 PRIOR FILING DATE: 1998-04-08  
77 PRIOR APPLICATION NUMBER: 60/081203  
78 PRIOR FILING DATE: 1998-04-09  
79 PRIOR APPLICATION NUMBER: 60/081229  
80 PRIOR FILING DATE: 1998-04-09  
81 PRIOR APPLICATION NUMBER: 60/081955  
82 PRIOR FILING DATE: 1998-04-15  
83 PRIOR APPLICATION NUMBER: 60/081817  
84 PRIOR FILING DATE: 1998-04-15  
85 PRIOR APPLICATION NUMBER: 60/081819  
86 PRIOR FILING DATE: 1998-04-15  
87 PRIOR APPLICATION NUMBER: 60/081952  
88 PRIOR FILING DATE: 1998-04-15  
89 PRIOR APPLICATION NUMBER: 60/081838  
90 PRIOR FILING DATE: 1998-04-15  
91 PRIOR APPLICATION NUMBER: 60/082568  
92 PRIOR FILING DATE: 1998-04-21  
93 PRIOR APPLICATION NUMBER: 60/082569  
94 PRIOR FILING DATE: 1998-04-21  
95 PRIOR APPLICATION NUMBER: 60/082704  
96 PRIOR FILING DATE: 1998-04-22  
97 PRIOR APPLICATION NUMBER: 60/082804  
98 PRIOR FILING DATE: 1998-04-22  
99 PRIOR APPLICATION NUMBER: 60/082700  
100 PRIOR FILING DATE: 1998-04-22  
101 PRIOR APPLICATION NUMBER: 60/082797  
102 PRIOR FILING DATE: 1998-04-22  
103 PRIOR APPLICATION NUMBER: 60/082796  
104 PRIOR FILING DATE: 1998-04-23  
105 PRIOR APPLICATION NUMBER: 60/083336  
106 PRIOR FILING DATE: 1998-04-27  
107 PRIOR APPLICATION NUMBER: 60/083322  
108 PRIOR FILING DATE: 1998-04-28  
109 PRIOR APPLICATION NUMBER: 60/083392  
110 PRIOR FILING DATE: 1998-04-29  
111 PRIOR APPLICATION NUMBER: 60/083495  
112 PRIOR FILING DATE: 1998-04-29  
113 PRIOR APPLICATION NUMBER: 60/083496  
114 PRIOR FILING DATE: 1998-04-29  
115 PRIOR APPLICATION NUMBER: 60/083499  
116 PRIOR FILING DATE: 1998-04-29  
117 PRIOR APPLICATION NUMBER: 60/083545  
118 PRIOR FILING DATE: 1998-04-29  
119 PRIOR APPLICATION NUMBER: 60/083554  
120 PRIOR FILING DATE: 1998-04-29  
121 PRIOR APPLICATION NUMBER: 60/083558  
122 PRIOR FILING DATE: 1998-04-29  
123 PRIOR APPLICATION NUMBER: 60/083559  
124 PRIOR FILING DATE: 1998-04-29  
125 PRIOR APPLICATION NUMBER: 60/083500  
126 PRIOR FILING DATE: 1998-04-29  
127 PRIOR APPLICATION NUMBER: 60/083742  
128 PRIOR FILING DATE: 1998-04-30  
129 PRIOR APPLICATION NUMBER: 60/084366  
130 PRIOR FILING DATE: 1998-05-05  
131 PRIOR APPLICATION NUMBER: 60/084414  
132 PRIOR FILING DATE: 1998-05-06  
133 PRIOR APPLICATION NUMBER: 60/084441  
134 PRIOR FILING DATE: 1998-05-06  
135 PRIOR APPLICATION NUMBER: 60/084637  
136 PRIOR FILING DATE: 1998-05-07  
137 PRIOR APPLICATION NUMBER: 60/084639  
138 PRIOR FILING DATE: 1998-05-07  
139 PRIOR APPLICATION NUMBER: 60/084640  
140 PRIOR FILING DATE: 1998-05-07  
141 PRIOR APPLICATION NUMBER: 60/084598  
142 PRIOR FILING DATE: 1998-05-07  
143 PRIOR APPLICATION NUMBER: 60/084600  
144 PRIOR FILING DATE: 1998-05-07  
145 PRIOR APPLICATION NUMBER: 60/084627  
146 PRIOR FILING DATE: 1998-05-07

```

; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match          0.7%  Score 8;  DB 9;  Length 67;
Best Local Similarity 100.0%; Pred. No. 36;
Matches      8;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1152 GLLLLALL 1159
Db      12 GLLLLALL 19

RESULT 68
US-09-999-832A-85
; Sequence 85, Application US/09999832A
; Publication No. US20020192706A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoli, Nicholas P.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PLC3
; CURRENT APPLICATION NUMBER: US/09/999,832A
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09

```

; PRIOR APPLICATION NUMBER: 60/081955  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081817  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081819  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081952  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081838  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/082568  
 ; PRIOR FILING DATE: 1998-04-21  
 ; PRIOR APPLICATION NUMBER: 60/082569  
 ; PRIOR FILING DATE: 1998-04-21  
 ; PRIOR APPLICATION NUMBER: 60/082704  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082804  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082700  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082797  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082796  
 ; PRIOR FILING DATE: 1998-04-23  
 ; PRIOR APPLICATION NUMBER: 60/083336  
 ; PRIOR FILING DATE: 1998-04-27  
 ; PRIOR APPLICATION NUMBER: 60/083322  
 ; PRIOR FILING DATE: 1998-04-28  
 ; PRIOR APPLICATION NUMBER: 60/083392  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083495  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083496  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083499  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083545  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083554  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083558  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083559  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083500  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083742  
 ; PRIOR FILING DATE: 1998-04-30  
 ; PRIOR APPLICATION NUMBER: 60/084366  
 ; PRIOR FILING DATE: 1998-05-05  
 ; PRIOR APPLICATION NUMBER: 60/084414  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: 60/084441  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: 60/084637  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084639  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084640  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084598  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084600  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084627  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084643  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/085339  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085338  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085323

; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085582  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085700  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085689  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085579  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085580  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085573  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085704  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 9; Length 67;  
 Best Local Similarity 100.0%; Pred. No. 36; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLALL 1159  
 Db 12 GLLLALL 19

RESULT 69

US-09-978-189-85  
 ; Sequence 85, Application US/09978189  
 ; Publication No. US2003004102A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Baker Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Baton, Dan  
 ; APPLICANT: Ferrara, Napoleon  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Hillan, Kenneth J.  
 ; APPLICANT: Kljavin, Ivar J.  
 ; APPLICANT: Kuo, Sophia S.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas F.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Shelton, David L.  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; FILE REFERENCE: P2630P1C7  
 ; CURRENT APPLICATION NUMBER: US/09/978,189  
 ; CURRENT FILING DATE: 2001-10-15  
 ; PRIOR APPLICATION NUMBER: 09/918585  
 ; PRIOR FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: 60/062250  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/064249  
 ; PRIOR FILING DATE: 1997-11-03  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066364  
 ; PRIOR FILING DATE: 1997-11-21



1 PRIOR APPLICATION NUMBER: 60/077450  
2 PRIOR FILING DATE: 1998-03-10  
3 PRIOR APPLICATION NUMBER: 60/077632  
4 PRIOR FILING DATE: 1998-03-11  
5 PRIOR APPLICATION NUMBER: 60/077641  
6 PRIOR FILING DATE: 1998-03-11  
7 PRIOR APPLICATION NUMBER: 60/077649  
8 PRIOR FILING DATE: 1998-03-11  
9 PRIOR APPLICATION NUMBER: 60/077791  
10 PRIOR FILING DATE: 1998-03-12  
11 PRIOR APPLICATION NUMBER: 60/078004  
12 PRIOR FILING DATE: 1998-03-13  
13 PRIOR APPLICATION NUMBER: 60/078886  
14 PRIOR FILING DATE: 1998-03-20  
15 PRIOR APPLICATION NUMBER: 60/078936  
16 PRIOR FILING DATE: 1998-03-20  
17 PRIOR APPLICATION NUMBER: 60/078910  
18 PRIOR FILING DATE: 1998-03-20  
19 PRIOR APPLICATION NUMBER: 60/078939  
20 PRIOR FILING DATE: 1998-03-20  
21 PRIOR APPLICATION NUMBER: 60/079294  
22 PRIOR FILING DATE: 1998-03-25  
23 PRIOR APPLICATION NUMBER: 60/079656  
24 PRIOR FILING DATE: 1998-03-26  
25 PRIOR APPLICATION NUMBER: 60/079664  
26 PRIOR FILING DATE: 1998-03-27  
27 PRIOR APPLICATION NUMBER: 60/079689  
28 PRIOR FILING DATE: 1998-03-27  
29 PRIOR APPLICATION NUMBER: 60/079663  
30 PRIOR FILING DATE: 1998-03-27  
31 PRIOR APPLICATION NUMBER: 60/079728  
32 PRIOR FILING DATE: 1998-03-27  
33 PRIOR APPLICATION NUMBER: 60/079786  
34 PRIOR FILING DATE: 1998-03-27  
35 PRIOR APPLICATION NUMBER: 60/079920  
36 PRIOR FILING DATE: 1998-03-30  
37 PRIOR APPLICATION NUMBER: 60/079923  
38 PRIOR FILING DATE: 1998-03-30  
39 PRIOR APPLICATION NUMBER: 60/080105  
40 PRIOR FILING DATE: 1998-03-31  
41 PRIOR APPLICATION NUMBER: 60/080107  
42 PRIOR FILING DATE: 1998-03-31  
43 PRIOR APPLICATION NUMBER: 60/080155  
44 PRIOR FILING DATE: 1998-03-31  
45 PRIOR APPLICATION NUMBER: 60/080194  
46 PRIOR FILING DATE: 1998-03-31  
47 PRIOR APPLICATION NUMBER: 60/080327  
48 PRIOR FILING DATE: 1998-04-01  
49 PRIOR APPLICATION NUMBER: 60/080328  
50 PRIOR FILING DATE: 1998-04-01  
51 PRIOR APPLICATION NUMBER: 60/080333  
52 PRIOR FILING DATE: 1998-04-01  
53 PRIOR APPLICATION NUMBER: 60/080334  
54 PRIOR FILING DATE: 1998-04-01  
55 PRIOR APPLICATION NUMBER: 60/081070  
56 PRIOR FILING DATE: 1998-04-08  
57 PRIOR APPLICATION NUMBER: 60/081049  
58 PRIOR FILING DATE: 1998-04-08  
59 PRIOR APPLICATION NUMBER: 60/081071  
60 PRIOR FILING DATE: 1998-04-08  
61 PRIOR APPLICATION NUMBER: 60/081195  
62 PRIOR FILING DATE: 1998-04-08  
63 PRIOR APPLICATION NUMBER: 60/081203  
64 PRIOR FILING DATE: 1998-04-09  
65 PRIOR APPLICATION NUMBER: 60/081229  
66 PRIOR FILING DATE: 1998-04-09  
67 PRIOR APPLICATION NUMBER: 60/081955  
68 PRIOR FILING DATE: 1998-04-15  
69 PRIOR APPLICATION NUMBER: 60/081817  
70 PRIOR FILING DATE: 1998-04-15  
71 PRIOR APPLICATION NUMBER: 60/081819  
72 PRIOR FILING DATE: 1998-04-15  
73 PRIOR APPLICATION NUMBER: 60/081952  
74 PRIOR FILING DATE: 1998-04-15  
75 PRIOR APPLICATION NUMBER: 60/081838  
76 PRIOR FILING DATE: 1998-04-15  
77 PRIOR APPLICATION NUMBER: 60/082568  
78 PRIOR FILING DATE: 1998-04-21  
79 PRIOR APPLICATION NUMBER: 60/082569  
80 PRIOR FILING DATE: 1998-04-21  
81 PRIOR APPLICATION NUMBER: 60/082704  
82 PRIOR FILING DATE: 1998-04-22  
83 PRIOR APPLICATION NUMBER: 60/082804  
84 PRIOR FILING DATE: 1998-04-22  
85 PRIOR APPLICATION NUMBER: 60/082700  
86 PRIOR FILING DATE: 1998-04-22  
87 PRIOR APPLICATION NUMBER: 60/082797  
88 PRIOR FILING DATE: 1998-04-22  
89 PRIOR APPLICATION NUMBER: 60/082796  
90 PRIOR FILING DATE: 1998-04-23  
91 PRIOR APPLICATION NUMBER: 60/083336  
92 PRIOR FILING DATE: 1998-04-27  
93 PRIOR APPLICATION NUMBER: 60/083322  
94 PRIOR FILING DATE: 1998-04-28  
95 PRIOR APPLICATION NUMBER: 60/083392  
96 PRIOR FILING DATE: 1998-04-29  
97 PRIOR APPLICATION NUMBER: 60/083495  
98 PRIOR FILING DATE: 1998-04-29  
99 PRIOR APPLICATION NUMBER: 60/083496  
100 PRIOR FILING DATE: 1998-04-29  
101 PRIOR APPLICATION NUMBER: 60/083499  
102 PRIOR FILING DATE: 1998-04-29  
103 PRIOR APPLICATION NUMBER: 60/083545  
104 PRIOR FILING DATE: 1998-04-29  
105 PRIOR APPLICATION NUMBER: 60/083554  
106 PRIOR FILING DATE: 1998-04-29  
107 PRIOR APPLICATION NUMBER: 60/083558  
108 PRIOR FILING DATE: 1998-04-29  
109 PRIOR APPLICATION NUMBER: 60/083559  
110 PRIOR FILING DATE: 1998-04-29  
111 PRIOR APPLICATION NUMBER: 60/083500  
112 PRIOR FILING DATE: 1998-04-29  
113 PRIOR APPLICATION NUMBER: 60/083742  
114 PRIOR FILING DATE: 1998-04-30  
115 PRIOR APPLICATION NUMBER: 60/084366  
116 PRIOR FILING DATE: 1998-05-05  
117 PRIOR APPLICATION NUMBER: 60/084414  
118 PRIOR FILING DATE: 1998-05-06  
119 PRIOR APPLICATION NUMBER: 60/084441  
120 PRIOR FILING DATE: 1998-05-06  
121 PRIOR APPLICATION NUMBER: 60/084637  
122 PRIOR FILING DATE: 1998-05-07  
123 PRIOR APPLICATION NUMBER: 60/084639  
124 PRIOR FILING DATE: 1998-05-07  
125 PRIOR APPLICATION NUMBER: 60/084640  
126 PRIOR FILING DATE: 1998-05-07  
127 PRIOR APPLICATION NUMBER: 60/084598  
128 PRIOR FILING DATE: 1998-05-07  
129 PRIOR APPLICATION NUMBER: 60/084600  
130 PRIOR FILING DATE: 1998-5-07  
131 PRIOR APPLICATION NUMBER: 60/084627  
132 PRIOR FILING DATE: 1998-05-07  
133 PRIOR APPLICATION NUMBER: 60/084643  
134 PRIOR FILING DATE: 1998-05-07  
135 PRIOR APPLICATION NUMBER: 60/085339  
136 PRIOR FILING DATE: 1998-05-13  
137 PRIOR APPLICATION NUMBER: 60/085338  
138 PRIOR FILING DATE: 1998-05-13  
139 PRIOR APPLICATION NUMBER: 60/085323  
140 PRIOR FILING DATE: 1998-05-13  
141 PRIOR APPLICATION NUMBER: 60/085582  
142 PRIOR FILING DATE: 1998-05-15  
143 PRIOR APPLICATION NUMBER: 60/085700  
144 PRIOR FILING DATE: 1998-05-15  
145 PRIOR APPLICATION NUMBER: 60/085689  
146 PRIOR FILING DATE: 1998-05-15

;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 10; Length 67;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159  
Db 12 GLLLLALL 19

## RESULT 70

US-09-978-608A-85  
; Sequence 85, Application US/09978608A  
; Publication No. US20030045462A1  
; GENERAL INFORMATION:

;; APPLICANT: Ashkenazi, Avi  
;; APPLICANT: Baker Kevin P.  
;; APPLICANT: Botstein, David  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Eaton, Dan  
;; APPLICANT: Ferrara, Napoleon  
;; APPLICANT: Filvaroff, Ellen  
;; APPLICANT: Fong, Sherman  
;; APPLICANT: Gerber, Hanspeter  
;; APPLICANT: Gerritsen, Mary E.  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Grimaldi, J. Christopher  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Hillan, Kenneth J.  
;; APPLICANT: Kijavin, Ivar J.  
;; APPLICANT: Kuo, Sophia S.  
;; APPLICANT: Napier, Mary A.  
;; APPLICANT: Pan, James;  
;; APPLICANT: Paoni, Nicholas F.  
;; APPLICANT: Roy, Margaret Ann  
;; APPLICANT: Shelton, David L.  
;; APPLICANT: Stewart, Timothy A.  
;; APPLICANT: Tumas, Daniel  
;; APPLICANT: Williams, P. Mickey  
;; APPLICANT: Wood, William I.  
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
;; FILE REFERENCE: P2630P1C22  
;; CURRENT APPLICATION NUMBER: US/09/978,608A  
;; CURRENT FILING DATE: 2001-10-16  
;; NUMBER OF SEQ ID NOS: 624  
;; Prior Application removed - See File Wrapper or Palm  
;; SEQ ID NO 85  
;; LENGTH: 67  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens

US-09-978-608A-85

Query Match 0.7%; Score 8; DB 10; Length 67;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159  
Db 12 GLLLLALL 19

## RESULT 71

US-09-978-585A-85  
; Sequence 85, Application US/09978585A  
; Publication No. US20030049633A1  
; GENERAL INFORMATION:

;; APPLICANT: Ashkenazi, Avi  
;; APPLICANT: Baker Kevin P.  
;; APPLICANT: Botstein, David  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Eaton, Dan  
;; APPLICANT: Ferrara, Napoleon  
;; APPLICANT: Filvaroff, Ellen  
;; APPLICANT: Fong, Sherman  
;; APPLICANT: Gao, Wei-Qiang  
;; APPLICANT: Gerber, Hanspeter  
;; APPLICANT: Gerritsen, Mary E.  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Grimaldi, J. Christopher  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Hillan, Kenneth J.  
;; APPLICANT: Kijavin, Ivar J.  
;; APPLICANT: Kuo, Sophia S.  
;; APPLICANT: Napier, Mary A.  
;; APPLICANT: Pan, James;  
;; APPLICANT: Paoni, Nicholas F.  
;; APPLICANT: Roy, Margaret Ann  
;; APPLICANT: Shelton, David L.  
;; APPLICANT: Stewart, Timothy A.  
;; APPLICANT: Tumas, Daniel  
;; APPLICANT: Williams, P. Mickey  
;; APPLICANT: Wood, William I.  
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
;; FILE REFERENCE: P2630P1C15  
;; CURRENT APPLICATION NUMBER: US/09/978,585A  
;; CURRENT FILING DATE: 2001-10-16  
;; NUMBER OF SEQ ID NOS: 624  
;; Prior Application removed - See File Wrapper or Palm  
;; SEQ ID NO 85  
;; LENGTH: 67  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens

US-09-978-585A-85

Query Match 0.7%; Score 8; DB 10; Length 67;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159  
Db 12 GLLLLALL 19

## RESULT 72

US-09-978-191A-85  
; Sequence 85, Application US/09978191A  
; Publication No. US20030050239A1  
; GENERAL INFORMATION:

;; APPLICANT: Ashkenazi, Avi  
;; APPLICANT: Baker Kevin P.  
;; APPLICANT: Botstein, David  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Eaton, Dan  
;; APPLICANT: Ferrara, Napoleon  
;; APPLICANT: Filvaroff, Ellen  
;; APPLICANT: Fong, Sherman  
;; APPLICANT: Gao, Wei-Qiang  
;; APPLICANT: Gerber, Hanspeter  
;; APPLICANT: Gerritsen, Mary E.  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James.  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630P1C4  
CURRENT APPLICATION NUMBER: US/09/978,191A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06

; PRIOR APPLICATION NUMBER: 60/084441  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: 60/084637  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084639  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084640  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084598  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084600  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084627  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084643  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/085339  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085338  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085323  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085582  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085700  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085689  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085579  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085580  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085573  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085704  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 10; Length 67;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159  
 DB 12 GLLLLALL 19

RESULT 73  
 US-09-978-403A-85  
 ; Sequence 85, Application US/09978403A  
 ; Publication No. US20030050240A1

; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Baker Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Ferrara, Napoleon  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Hillan, Kenneth J.  
 ; APPLICANT: Kljavin, Ivar J.  
 ; APPLICANT: Kuo, Sophia S.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James;  
 ; APPLICANT: Paoni, Nicholas P.

; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Shelton, David L.  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; FILE OF INVENTION: Acids Encoding the Same  
 ; FILE REFERENCE: P2630P1C17  
 ; CURRENT APPLICATION NUMBER: US/09/978,403A  
 ; CURRENT FILING DATE: 2002-03-19  
 ; PRIOR APPLICATION NUMBER: 09/918585  
 ; PRIOR FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: 60/062250  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/064249  
 ; PRIOR FILING DATE: 1997-11-03  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066364  
 ; PRIOR FILING DATE: 1997-11-21  
 ; PRIOR APPLICATION NUMBER: 60/077450  
 ; PRIOR FILING DATE: 1998-03-10  
 ; PRIOR APPLICATION NUMBER: 60/077632  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077641  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077649  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077791  
 ; PRIOR FILING DATE: 1998-03-12  
 ; PRIOR APPLICATION NUMBER: 60/078004  
 ; PRIOR FILING DATE: 1998-03-13  
 ; PRIOR APPLICATION NUMBER: 60/078886  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078936  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078939  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/079294  
 ; PRIOR FILING DATE: 1998-03-25  
 ; PRIOR APPLICATION NUMBER: 60/079656  
 ; PRIOR FILING DATE: 1998-03-26  
 ; PRIOR APPLICATION NUMBER: 60/079664  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079689  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079663  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079728  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079786  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079920  
 ; PRIOR FILING DATE: 1998-03-30  
 ; PRIOR APPLICATION NUMBER: 60/079923  
 ; PRIOR FILING DATE: 1998-03-30  
 ; PRIOR APPLICATION NUMBER: 60/080105  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080107  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080165  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080194  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080327  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080328  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080333  
 ; PRIOR FILING DATE: 1998-04-01

;; PRIOR APPLICATION NUMBER: 60/080334  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/081070  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081049  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081071  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081195  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081955  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081819  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081952  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081838  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082568  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082569  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082704  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082804  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082700  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082797  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082796  
;; PRIOR FILING DATE: 1998-04-23  
;; PRIOR APPLICATION NUMBER: 60/083336  
;; PRIOR FILING DATE: 1998-04-27  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083392  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083495  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083496  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083499  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083554  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083558  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083559  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083500  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083742  
;; PRIOR FILING DATE: 1998-04-30  
;; PRIOR APPLICATION NUMBER: 60/084366  
;; PRIOR FILING DATE: 1998-05-05  
;; PRIOR APPLICATION NUMBER: 60/084414  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084441  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084639  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084640

;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084598  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 10; Length 67;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLJLALL 1159  
|||  
DB 12 GLLJLALL 19

RESULT 74  
US-09-978-564A-85  
; Sequence 85, Application US/09978564A  
; Publication No. US20030050241A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Geritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kllavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

1 TITLE OF INVENTION: Acids Encoding the Same  
2 FILE REFERENCE: P2630PIC25  
3 CURRENT APPLICATION NUMBER: US/09/978,564A  
4 CURRENT FILING DATE: 2001-10-16  
5 PRIOR APPLICATION NUMBER: 09/918585  
6 PRIOR FILING DATE: 2001-07-30  
7 PRIOR APPLICATION NUMBER: 60/062250  
8 PRIOR FILING DATE: 1997-10-17  
9 PRIOR APPLICATION NUMBER: 60/064249  
10 PRIOR FILING DATE: 1997-11-03  
11 PRIOR APPLICATION NUMBER: 60/065311  
12 PRIOR FILING DATE: 1997-11-13  
13 PRIOR APPLICATION NUMBER: 60/066364  
14 PRIOR FILING DATE: 1997-11-21  
15 PRIOR APPLICATION NUMBER: 60/077450  
16 PRIOR FILING DATE: 1998-03-10  
17 PRIOR APPLICATION NUMBER: 60/077632  
18 PRIOR FILING DATE: 1998-03-11  
19 PRIOR APPLICATION NUMBER: 60/077641  
20 PRIOR FILING DATE: 1998-03-11  
21 PRIOR APPLICATION NUMBER: 60/077649  
22 PRIOR FILING DATE: 1998-03-11  
23 PRIOR APPLICATION NUMBER: 60/077791  
24 PRIOR FILING DATE: 1998-03-12  
25 PRIOR APPLICATION NUMBER: 60/078004  
26 PRIOR FILING DATE: 1998-03-13  
27 PRIOR APPLICATION NUMBER: 60/078886  
28 PRIOR FILING DATE: 1998-03-20  
29 PRIOR APPLICATION NUMBER: 60/078936  
30 PRIOR FILING DATE: 1998-03-20  
31 PRIOR APPLICATION NUMBER: 60/078910  
32 PRIOR FILING DATE: 1998-03-20  
33 PRIOR APPLICATION NUMBER: 60/078939  
34 PRIOR FILING DATE: 1998-03-20  
35 PRIOR APPLICATION NUMBER: 60/079294  
36 PRIOR FILING DATE: 1998-03-25  
37 PRIOR APPLICATION NUMBER: 60/079656  
38 PRIOR FILING DATE: 1998-03-26  
39 PRIOR APPLICATION NUMBER: 60/079664  
40 PRIOR FILING DATE: 1998-03-27  
41 PRIOR APPLICATION NUMBER: 60/079689  
42 PRIOR FILING DATE: 1998-03-27  
43 PRIOR APPLICATION NUMBER: 60/079663  
44 PRIOR FILING DATE: 1998-03-27  
45 PRIOR APPLICATION NUMBER: 60/079728  
46 PRIOR FILING DATE: 1998-03-27  
47 PRIOR APPLICATION NUMBER: 60/079786  
48 PRIOR FILING DATE: 1998-03-27  
49 PRIOR APPLICATION NUMBER: 60/079920  
50 PRIOR FILING DATE: 1998-03-30  
51 PRIOR APPLICATION NUMBER: 60/079923  
52 PRIOR FILING DATE: 1998-03-30  
53 PRIOR APPLICATION NUMBER: 60/080105  
54 PRIOR FILING DATE: 1998-03-31  
55 PRIOR APPLICATION NUMBER: 60/080107  
56 PRIOR FILING DATE: 1998-03-31  
57 PRIOR APPLICATION NUMBER: 60/080165  
58 PRIOR FILING DATE: 1998-03-31  
59 PRIOR APPLICATION NUMBER: 60/080194  
60 PRIOR FILING DATE: 1998-03-31  
61 PRIOR APPLICATION NUMBER: 60/080327  
62 PRIOR FILING DATE: 1998-04-01  
63 PRIOR APPLICATION NUMBER: 60/080328  
64 PRIOR FILING DATE: 1998-04-01  
65 PRIOR APPLICATION NUMBER: 60/080333  
66 PRIOR FILING DATE: 1998-04-01  
67 PRIOR APPLICATION NUMBER: 60/080334  
68 PRIOR FILING DATE: 1998-04-01  
69 PRIOR APPLICATION NUMBER: 60/081070  
70 PRIOR FILING DATE: 1998-04-08  
71 PRIOR APPLICATION NUMBER: 60/081049  
72 PRIOR FILING DATE: 1998-04-08  
73 PRIOR APPLICATION NUMBER: 60/081071  
74 PRIOR FILING DATE: 1998-04-08  
75 PRIOR APPLICATION NUMBER: 60/081195  
76 PRIOR FILING DATE: 1998-04-08  
77 PRIOR APPLICATION NUMBER: 60/081203  
78 PRIOR FILING DATE: 1998-04-09  
79 PRIOR APPLICATION NUMBER: 60/081229  
80 PRIOR FILING DATE: 1998-04-09  
81 PRIOR APPLICATION NUMBER: 60/081955  
82 PRIOR FILING DATE: 1998-04-15  
83 PRIOR APPLICATION NUMBER: 60/081817  
84 PRIOR FILING DATE: 1998-04-15  
85 PRIOR APPLICATION NUMBER: 60/081819  
86 PRIOR FILING DATE: 1998-04-15  
87 PRIOR APPLICATION NUMBER: 60/081952  
88 PRIOR FILING DATE: 1998-04-15  
89 PRIOR APPLICATION NUMBER: 60/081838  
90 PRIOR FILING DATE: 1998-04-15  
91 PRIOR APPLICATION NUMBER: 60/082568  
92 PRIOR FILING DATE: 1998-04-21  
93 PRIOR APPLICATION NUMBER: 60/082569  
94 PRIOR FILING DATE: 1998-04-21  
95 PRIOR APPLICATION NUMBER: 60/082704  
96 PRIOR FILING DATE: 1998-04-22  
97 PRIOR APPLICATION NUMBER: 60/082804  
98 PRIOR FILING DATE: 1998-04-22  
99 PRIOR APPLICATION NUMBER: 60/082700  
100 PRIOR FILING DATE: 1998-04-22  
101 PRIOR APPLICATION NUMBER: 60/082797  
102 PRIOR FILING DATE: 1998-04-22  
103 PRIOR APPLICATION NUMBER: 60/082796  
104 PRIOR FILING DATE: 1998-04-23  
105 PRIOR APPLICATION NUMBER: 60/083336  
106 PRIOR FILING DATE: 1998-04-27  
107 PRIOR APPLICATION NUMBER: 60/083322  
108 PRIOR FILING DATE: 1998-04-28  
109 PRIOR APPLICATION NUMBER: 60/083392  
110 PRIOR FILING DATE: 1998-04-29  
111 PRIOR APPLICATION NUMBER: 60/083495  
112 PRIOR FILING DATE: 1998-04-29  
113 PRIOR APPLICATION NUMBER: 60/083496  
114 PRIOR FILING DATE: 1998-04-29  
115 PRIOR APPLICATION NUMBER: 60/083499  
116 PRIOR FILING DATE: 1998-04-29  
117 PRIOR APPLICATION NUMBER: 60/083545  
118 PRIOR FILING DATE: 1998-04-29  
119 PRIOR APPLICATION NUMBER: 60/083554  
120 PRIOR FILING DATE: 1998-04-29  
121 PRIOR APPLICATION NUMBER: 60/083558  
122 PRIOR FILING DATE: 1998-04-29  
123 PRIOR APPLICATION NUMBER: 60/083559  
124 PRIOR FILING DATE: 1998-04-29  
125 PRIOR APPLICATION NUMBER: 60/083500  
126 PRIOR FILING DATE: 1998-04-29  
127 PRIOR APPLICATION NUMBER: 60/083742  
128 PRIOR FILING DATE: 1998-04-30  
129 PRIOR APPLICATION NUMBER: 60/084366  
130 PRIOR FILING DATE: 1998-05-05  
131 PRIOR APPLICATION NUMBER: 60/084414  
132 PRIOR FILING DATE: 1998-05-06  
133 PRIOR APPLICATION NUMBER: 60/084441  
134 PRIOR FILING DATE: 1998-05-06  
135 PRIOR APPLICATION NUMBER: 60/084637  
136 PRIOR FILING DATE: 1998-05-07  
137 PRIOR APPLICATION NUMBER: 60/084639  
138 PRIOR FILING DATE: 1998-05-07  
139 PRIOR APPLICATION NUMBER: 60/084640  
140 PRIOR FILING DATE: 1998-05-07  
141 PRIOR APPLICATION NUMBER: 60/084598  
142 PRIOR FILING DATE: 1998-05-07  
143 PRIOR APPLICATION NUMBER: 60/084600  
144 PRIOR FILING DATE: 1998-05-07  
145 PRIOR APPLICATION NUMBER: 60/084627  
146 PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/084643  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/085339  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085338  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085323  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085582  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085700  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085689  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085579  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085580  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085573  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085704  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 10; Length 67;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLALL 1159

Db 12 GLLALL 19

RESULT 75

US-09-999-833A-85  
 ; Sequence 85, Application US/09999833A  
 ; Publication No. US20030054405A1  
 ; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Baker Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Baton, Dan  
 ; APPLICANT: Ferrara Napoleon  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Hillan, Kenneth J.  
 ; APPLICANT: Kljavin, Ivar J.  
 ; APPLICANT: Kuo, Sophia S.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas F.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Shelton, David L.  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; FILE OF INVENTION: Acids Encoding the Same

FILE REFERENCE: P26301C65

CURRENT APPLICATION NUMBER: US/09/999,833A

CURRENT FILING DATE: 2001-10-24

PRIOR APPLICATION NUMBER: 09/918595

PRIOR FILING DATE: 2001-07-30

PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/064249  
 ; PRIOR FILING DATE: 1997-11-03  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066364  
 ; PRIOR FILING DATE: 1997-11-21  
 ; PRIOR APPLICATION NUMBER: 60/077450  
 ; PRIOR FILING DATE: 1998-03-10  
 ; PRIOR APPLICATION NUMBER: 60/077632  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077641  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077649  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077791  
 ; PRIOR FILING DATE: 1998-03-12  
 ; PRIOR APPLICATION NUMBER: 60/078004  
 ; PRIOR FILING DATE: 1998-03-13  
 ; PRIOR APPLICATION NUMBER: 60/078886  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078936  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078939  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/079294  
 ; PRIOR FILING DATE: 1998-03-25  
 ; PRIOR APPLICATION NUMBER: 60/079656  
 ; PRIOR FILING DATE: 1998-03-26  
 ; PRIOR APPLICATION NUMBER: 60/079664  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079689  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079663  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079728  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079786  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079920  
 ; PRIOR FILING DATE: 1998-03-30  
 ; PRIOR APPLICATION NUMBER: 60/079923  
 ; PRIOR FILING DATE: 1998-03-30  
 ; PRIOR APPLICATION NUMBER: 60/080105  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080107  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080165  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080194  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080327  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080328  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080333  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080334  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/081070  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081049  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081071  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081195  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081203  
 ; PRIOR FILING DATE: 1998-04-09  
 ; PRIOR APPLICATION NUMBER: 60/081229  
 ; PRIOR FILING DATE: 1998-04-09



```

; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323

```

```

; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

```

```

Query Match          0.7%; Score 8; DB 10; Length 67;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1152 GLLHLL 1159
Db 12 GLLHLL 19

```

```

Search completed: June 24, 2004, 18:11:59
Job time : 58 secs

```